1   **Assertion, presumption and presupposition:**

2   **an account of the erstwhile nominalizer YUM in Khalkha Mongolian**[i]

3

4   Benjamin Brosig[1,2], Foong Ha Yap[1,3], Kathleen Ahrens[1] [ii]

5   *Hong Kong Polytechnic University*[1]

6   *Academia Sinica*[2]

7   *Chinese University of Hong Kong, Shenzhen*[3]

8

9

10   *In memory of John C. Street (1930-2017)*

11

12   **Abstract:**

13

14   In this paper, we analyze the clitic YUM (< 'thing') in Khalkha Mongolian which, in different

15   syntactic contexts, reinforces assertiveness or expresses different shades of presumption or

16   presupposition. The former holds for declaratives where the presence of YUM conveys the

17   speaker's strong subjective commitment. In question clauses, YUM is used to indicate the

18   speaker's subjective and often strong guess, sometimes to the point that the speaker

19   presupposes that the proposition actually obtains. In subordinate clauses, YUM may fulfill

20   the same function or serve as a structurally necessary nominalizer for adjectival predicates

21   without introducing any semantic opposition. In declaratives marked as immediately

22   perceived, YUM conveys inference via assumptive reasoning. We thus analyze YUM as the

23   marker of subjective speaker conviction that within the Khalkha Mongolian declarative

24   system is opposed to both simple factuality and overt evidential marking.

25

26   **Keywords:** assertion; presumption; presupposition; inference; evidentiality

27

28

29   **1. Introduction**

30

31   In studies related to language philosophy and pragmatics, the question has been raised as to what

32   it means to make and interpret an assertion. On the one hand, it has been assumed that the

33   interpretation of assertions entails that the hearer attributes a belief to the speaker. Interpreting

34   would thus be the reverse of asserting in some Gricean approaches (e.g. Bach & Harnish 1979:

35   16), i.e. making a statement to the effect that a given proposition holds true, for which the speaker

36   must intend the hearer to think that the speaker believes said proposition. Conversely, hearers

37   may conceive of assertions as primary information similar to information perceived through

38   observation (Dummett 1981: 355), and speakers, in order to assert, merely need to intend to

39   impart information about the world. The latter understanding is supported by the cognitive

40   development of children who make and interpret assertions before they can attribute beliefs.

41   Fully competent speakers, in turn, may choose to interpret a given statement as the speaker's

42   mere belief rather than immediate information about reality e.g. if they have reason to doubt its

43   actual veracity or are participating in a type of discourse (such as scientific discourse) that entails

44   the intent to persuade rather than to communicate facts about the world as true (cf. Jary 2010:

45   37-44).

46   In the above-mentioned disciplines, the meaning of assertions is mostly presumed to be

47 language-independent (cf. Pagin 2016). However, since the realization of assertions is sensitive to

48 discourse type, one would expect that it may also vary between languages. For instance, the

49 Siouan language Hidatsa (Matthews 1965: 99-100; also quoted in Palmer 2001: 68-69)

50 distinguishes two types of assertion markers. The weaker marker indicates that the speaker

51 BELIEVES the statement to be true and would be mistaken if it turned out to be wrong. But by using

52 the stronger marker, the speaker claims that s/he KNOWS the statement to be true and could

53 otherwise be considered a liar. The starting point, then, could be that speech acts often entail

54 particular relationships between interlocutors that differ cross-culturally. Consider, for example,

55 the speech act encoded by English *warn* or *threaten,* compared to that for Japanese *satos-* which

56 invokes a benevolent, responsible, higher-ranking speaker. Social principles for interpreting

57 linguistic constructions such as modal questions may likewise vary cross-culturally (Wierzbicka

58 2003: 152-154). Now if constructions exist in hierarchical networks, then even very general

59 concepts such as subject and object are best understood as language-specific semantically related

60 groupings of participant roles of events in conceptual space (cf. Goldberg 1995, Croft 2001: 132-

61 171, Barðdal 2011). A similar case could then be made for the language-specific semantics of plain

62 declarative constructions due to differing conversational conventions, and this is indeed what

63 Behrens (2012) sets out to do for English and Hungarian. Plain English declaratives appear to

64 require hedging (such as *I think*) if they are used to express subjective opinions, while their

65 Hungarian equivalents do not (and structurally similar linguistic devices like *én azt hiszem* are

66 rather used to reinforce beliefs). Plain English declaratives would thus be closer to strong

67 assertives in Hidatsa, while plain Hungarian declaratives would resemble its weak assertives.

68   While English or Hungarian might each (possibly) field a single declarative construction with

69 language-internally consistent epistemological meaning, things become more complicated in

70 languages with grammaticalized evidentiality. Evidentials differ from modals in that they don't

71 necessarily relativize the speaker's own claim to factuality, but communicate the speaker's

72 knowledge (cf. Aikhenvald & Dixon [eds.] 2014's *Grammar of knowledge*) given the way how a

73 speaker has accessed or learnt about it (Tournadre & LaPolla 2014). Direct sensory access to an

74 event does not necessarily imply reliable information (Lazard 1999 on Persian, Johanson 2006 on

75 Turkish, Brosig 2014 on Middle Mongol) and thus may well be used (in conventions established

76 by specific language communities) if a speaker merely wants to convey her sensory perception of

77 an event that she, in terms of world knowledge, is not entirely convinced of (cf. Jary [2010: 40]'s

78 label "perception-by-proxy").

79   One such system that operates with a number of epistemologically qualifying categories is

80 found in Khalkha Mongolian. Here, most indicative (declarative or interrogative) forms distinguish

81 categories such as tense, time of acquisition, evidentiality and, marginally, surprise. First, most

82 indicative (declarative or interrogative) forms distinguish between immediate and established

83 knowledge.[1] Forms that code immediate knowledge (*-laa*, *-žee*, *[bai]-n*) are morphologically finite

84 (i.e. can only be used in finite position) and refer to events that the speaker realizes, re-accesses,

85 evaluates or re-evaluates at the time of speech. In the past tense, they are subdivided according

86 to evidentiality into the type of personal access that the speaker had to the event, i.e. whether

87 the speaker perceived the event directly (*-laa*) or inferred it from directly perceivable evidence (-

88 *žee*). Established forms (*-saŋ*, *-dag*, [*bai[ɢ]]-aa*), in contrast, are morphologically participles in finite

89 uses. They refer to information has already been integrated as facts into the speaker's mental

90 representation of the world (cf. Brosig 2015, 2018, Brosig & Skribnik 2018).

---

[1] This excludes imperatives and hortatives that are expressed by suffixes that would occupy the same slot (see Janhunen 2012). It also excludes the suffix *-n* which refers to temporally unspecific potential developments (Brosig 2015: 51-63) when it is attached to verbs other than *bai-* (see next paragraph). It is unclear whether it extends to surprise-related uses of the habitual participle in *-dag* (Brosig 2015: 93-96, 101-104) and to the old past marker *-w* (Brosig 2018: 61-62).

91   Since past tense is always marked overtly, speakers always have to distinguish between

92   established (-*saŋ*), direct (-*laa*) and inferred (-*žee*) knowledge. The distinctions in the present

93   tense are more complicated: evidentiality is not expressed, but time of acquisition closely

94   interacts with the aspectuality. For temporally delimted (stage-level) predications, speakers

95   always have to distinguish between events that they actively perceive (via the senses or in their

96   mind) or infer at the time of speech (marked by *X + bai-n*) as in (1a) and those that they have

97   already been aware of for a while and for which no currently available evidence is resorted to

98   (marked by *X + bai-[ɢ]aa*) as in (1b). Regular verbal habitual predications (marked by -*dag*) and

99   nominal predications that denote properties of an individual (marked by bare nominal predicates)

100 usually lack explicit marking for time of realization, as in (1c), though this type of marking

101 ultimately tends to code established knowledge. However, these dividing lines are not carved in

102 stone. Habitual markers are occasionally used within the scope of an explicit marker of time of

103 realization such as *bai-n* in (1d),[2] and habitually re-occurring events with an internal time structure

104 may use a complex aspect-marking construction in the scope of habitual marking (*X + bai-dag*) as

105 in (1e).[3]

106

107 (1) (a) *nasn-ii xan'-aa*   ***xai-ž   bai-n***. (IC)

108   life-GEN longterm.sexual.partner-RPOSS search-CVB AUX-PRS.IMM

---

[2] In a Spoken Corpus (SC, see §2) which is biased towards autobibliographical narrative data, bare habitual forms were used in 98,4% (n=479) of all instances, while established marking in -*dag bai-(ɢ)aa* was used in 7 cases and immediate marking in -*dag bai-n* in 1 case (see Brosig 2015). In the more balanced Internet Corpus (IC), the immediate perception form -*dag bai-n* is 20,3 times as frequent (10844:535) as the form for established perception -*dag bai-[ɢ]aa*. Notably, inference in scientific and administrative texts is very often marked by *X + bai-n*.

[3] "X" here is used to represent three different types of dependent aspectual structures (corresponding to the three rightmost columns of Table 1). First, it represents progressive, continuative-resultative and perfect marking via the converbal suffixes -*ž* and -*aad* and the participial suffix -*saŋ*, respectively. These dependent aspectual constructions, followed by the copular auxiliary *bai-* as a dummy verb, can combine with all finite and finitely used markers. The dependent marking of habituality by -*dag bai-*, on the other hand, cannot combine with finitely used -*dag*. Finally, if the main predicate is not a verb but a nominal, it does not require any additional marking if no temporal delimitation is intended.

109   '[I] <u>am searching</u> for my life partner {<u>consciously, as I am telling you</u>}.'

110   (b) *min-ii üyiinxeŋ saiŋ nöxör **xai-ž bai-(ɢ)aa***. (IC)

111   1SG-GEN peer   good husband search-CVB AUX-PRS.EST

112   'My friend <u>is searching [without me being present]</u> for a good husband.'

113   (c) *Iim üy-d az.žargal-tai bai-x öör xüŋ **xai-dag***. (IC)

114   such time-DAT happiness-COM AUX-PTCP.NPST other person search-PTCP.HAB

115   ([Men] wait for women to make the first step. If the woman doesn't make the first

116   step, they lose confidence in themselves or even think that the woman is rejecting

117   them.) 'At such a point, [men] <u>search</u> for another person to be happy with.'

118   (d) *Üün-tei xolboo-toi-goor er-čüüd öör xeŋ.negeŋ seks-iin*

119   DEM.PROX-COM connection-COM-INS male-PL other somebody sex-GEN

120   *xamtragč emegtei-g **er-ž xai-dag bai-n***. (IC)

121   partner   woman-ACC search-CVB search-PTCP.HAB AUX-PRS.IMM

122   'In connection with this [<u>e.g. as I have observed on multiple occasions</u>], men <u>look</u>

123   for some other sex partner.'

124   (e) *bid ürgelž xeŋ.negn-iig **xai-ž bai-dag***. (IC)

125   1PL permanently somebody-ACC search-CVB AUX-PTCP.HAB

126   'We'<u>re always [in regular intervals] [in the course of] searching</u> for somebody.'

127

128   Table 1 gives an overview of this system for past and present tense forms.[4]

129

---

[4] In the Khalkha Mongolian past tense, the same aspectual distinctions are found as in the present, but since they don't seem to influence aspectual choices (in contrast to Middle Mongol [Brosig&Skribnik 2018: 557-559] and Oirat [Bläsing 1984: 28, Gotō 2009: 127-8]). In the table, the past tense forms are not listed individually, except in the case of habitually re-occuring past forms, which we included into the past tense column to facilitate comparison between past and present habitual forms. Future reference is achieved via potential *-n*, present tense prospectives, perfective past suffixes and actional markers in an eclectic system that will have to be described in a future paper.

130   *[intended location of Table 1]*

131

132   Since immediate perceptual forms are also used for visual and suppostive contexts, while

133   established markers entail a certain reification and abstraction from concrete memory (see Brosig

134   2018), there is no reason to assume that they are ranked in terms of the speaker's commitment.

135   The evidential and time-of-acquisition-related distinctions are primarily intended to enable the

136   addressee to properly assess the truth of an assertion, thus relegating part of the responsibility

137   for accepting it to the addressee.

138   However, there is one additional distinction shown in Table 1 where a clitic =*iin* ~ =*n* (or a

139   particle *yum*) can be attached to participles or other nominals under certain circumstances, either

140   in finite position or before the copular auxiliary *bai*-. This particle originated from an abstract noun

141   *yum* (*jaʊma) that can refer to physical entities as in (2a) and events as in (2b).[5] As a clitic in post-

142   predicative position, it can reinforce an assertion as in (2c) or mark a proposition as presumed or

143   presupposed as in (2d), while it appears to somewhat lower the degree of epistemic certainty

144   when expressing presumptive reasoning as in (2e).[6] As an erstwhile functional noun, =*iin* has

145   retained most of the distributional properties of nominal *yum* in that it is regularly "modified" by

146   participles and adjectives and cannot be preceded by finite verbs. It has, however, developed a

147   limited ability to cliticize to nouns.

148

---

[5] While reflexes of *jaʊma~*jama are common to all modern Mongolic languages (Nugteren 2011: 544-545), it is first attested in a single Late Western Middle Mongol source and thus cannot be reconstructed for Proto-Mongolic. It might have developed from either Early Middle Mongol (13th century) *yau ba* what even 'anything' (John Street, p.c., 2017-01-29) or an unattested Late Middle Mongol form **yau ma* 'anything' < *yau* 'what' + Turkic *yämä* 'also' (Hans Nugteren 2011: 544, p.c., 2017-02-15).

[6] In this paper, we define the term PRESUPPOSE from a speaker's vantage point as "for the speaker to presume a proposition to hold and to further assume that the addressee can plausibly be expected to accept the proposition as a background assumption as well". At the same time, we define PRESUME as "for the speaker to assume a proposition to hold on the basis of defeasible reasoning" (both definitions partially adapted from Macagno 2018: 299-308).

149   (2) (a) *ter*   **yum**   *id-seŋ*.

150   that   thing   eat-PST.EST

151   'She ate <u>something</u>.'

152   (b) *iim*   **yum**   *tox'oold-soŋ*.

153   such   thing   happen-PST.EST

154   'Such <u>an event</u> happened.'

155   (c) *ter*   *id-s=**iin***.

156   that   eat-PTCP.PRF=ASS

157   'She ate<u>!</u>'

158   (d) *ter*   *id-s=**iim***   **bol**   *saiŋ*.

159   that   eat-PTCP.PRF=ASS   if   good

160   'If she has <u>indeed</u> eaten, that's good.'

161   (e) *ter*   *id-s=**iim***   **bai-n.**

162   that   eat-PTCP.PRF=ASS/NMLZ AUX-PRS.IMM

163   'She has eaten / <u>must</u> have eaten.' (inferred at speech time)

164

165   In studies on the predicative or past tense system of Khalkha Mongolian (e.g. Byambasaŋ et al.

166   1987, Svantesson 1991, Song 1997, Binnick 2012, Brosig 2015a, 2018), the function of *yum* has

167   not been integrated into the analysis since it would be regarded as a "sentence-final particle" and

168   thus be assigned to another structural domain. Conversely, Kang (2003: 28-33)'s study of

169   sentence-final particles features *yum* but focuses on distribution rather than meaning. There are,

170   however, three dedicated studies on the particle *yum*. Möŋx-Amɢalaŋ (1996) investigates *yum* in

171   collocations with interrogative words, participles and other nominals, sentence-final particles,

172   copulas, conjunctions and negators. He captures the respective meanings of these collocations by

173   using Mongolian speech act verbs (such as *xüürneŋ batalsaŋ* 'asserted' or *čuxalčlaŋ sonsɢosoŋ*

174 'emphasized as important' for the collocation -*saŋ yum*) and sometimes in terms of modal

175 categories. Mukai (2001) examines *yum* at the end of declarative main clauses (positive or

176 negated by *biš*) and interrogatives. Primarily focusing on information structure, Mukai analyzes

177 the 'focus adjustment form' 焦点調整形式 *yum* as a nominalizer that shifts attention away from

178 the main predicate to some other sentence constituent that is either syntactically peripheral or

179 semantically detailed or defined by contrast to adjacent clauses. He lists several uses not covered

180 by his analysis, but suggests that they are secondary extensions. Jīngāng (2007) also examines at

181 the end of declarative constructions and, building on Mukai's discussion of *yum* as a nominalizer,

182 claims a development from a general noun for physical entities 名詞 (もの) via an event-referring

183 noun こと（状能）to a noun that refers to 'generally known facts' 一般通念的こと, and from

184 there either to a 'result that necessarily follows from a causal relation' 因果関係による必然的

185 結果 or a 'reason why something occurred' 事柄成立の理由.

186   What these studies do not yet take into the picture is how the semantic contribution of *yum*

187 in subordinate clauses relates to its meaning in matrix clauses. Möŋx-Amɢalaŋ doesn't adddress

188 this problem because he is not concerned with the distinction between meaning and contextual

189 implicatures or the semantic contribution of *yum* to the collocations it partakes in. Mukai and

190 Jīngāng don't discuss these types of structures at all, emphasizing meanings or connotations that

191 are restricted to some of the attested contexts. In this paper, then, we set out to describe the

192 functions that *yum* fulfills in several syntactic positions in an attempt to come up with a unified

193 account of its semantic contribution to declarative, interrogative and subordinate clauses that

194 accounts for its assertive meaning in clause-final position and its presumptive to presuppositional

195 meaning in non-final position.

196   To set the foundation of our study, we will first provide preliminary information about our

197   linguistic data, methods and transcription conventions (§2). Next we will provide an overview of

198   the construction types that *yum* is involved in (§3), and discuss its morphological and morphemic

199   status (§4). Following that, we will discuss *yum* as a referential noun (§5). We then discuss the

200   assertive and presumptive functions of *yum* in sentence-final position (§6), in post-predicative

201   position as part of clause connectors (§7), and in the presumptive reasoning construction (§8). In

202   the conclusion (§9), we will summarize our findings and explore the implications that

203   grammaticalized *yum* has for the overall system of assertive distinctions encoded in the Khalkha

204   Mongolian tense-aspect-evidentiality system.

205

206   **2. Data, methods and conventions**

207

208   This study takes as its starting point a spoken corpus (SC) of Khalkha Mongolian (Zolžargal & Brosig

209   2012) consisting of 60,000 words mostly taken from unscripted TV programs (see Brosig 2015a).

210   All tokens found in this corpus were categorized by function and analyzed within their discourse

211   context. [7] When conducting this analysis, we used Brosig (2015a, 2018)'s fieldnotes from

212   contrastive elicitation and the classification of Mönx-Amɢalaŋ (1996) as our starting points. We

213   relied on SC to identify *yum*-based construction types and analyze constructions with a high token

214   frequency. We additionally relied on a larger internet corpus (IC) consisting of 34.6 million words

215   in Cyrillic script (Östling & Brosig 2011) to clarify the use of some patterns with low token

216   frequency. Where relevant, we also cite Mönx-Amɢalaŋ (1996)'s examples which were taken from

---

[7] For identifying tokens in SC, we partially relied on the tagging used by Brosig (2015a, 2018). Here, the verbs *teg-* 'do like that', *iŋg-* 'do like this' and *yaa-* 'do how' were excluded if not accompanied by overt arguments, the verb *bai-* was only annotated as an auxiliary but not as a locational/possessive verb, and the verbs *bol-* 'become' and *ge-* 'say' (a quotative verb) were mostly excluded when not preceded by verbs.

217   a corpus of nine theatre plays. For retrieving specific lexical collocations for expositional rather

218   than heuristic purposes, we occasionally used the internet search engine Google. The source of

219   each individual example is indicated throughout the text. Elicitation involving 2-4 informants was

220   used to check the results of text analysis and distributional analysis throughout most of the paper.

221   It played no role for most of the structure-oriented data in §5, while the distributionally

222   contradictory data in §7.2 was discussed with 6-8 informants. Informants consisted of a

223   convenience sample of adults between 18 and 50 years of age who were living in Ulaanbaatar,

224   had not studied Mongolian at university level nor worked as teachers, and had stated that both

225   of their parents were Khalkha speakers.

226   The language variety investigated in this study is the Khalkha dialect of Mongolian, the

227   dominant language of the Mongolian state. Findings would presumably also hold for other Central

228   Mongolian dialects such as Chakhar and Shilingol, but we did not undertake to demonstrate this.

229   The distribution of *yum* in other branches of Central Mongolic such as Eastern Mongolian and

230   Western Mongolian (=Oirat) is briefly mentioned for two dialects, Khorchin and Kalmyk, which

231   might or might not be representative of their branches. For instance, there is cursory evidence

232   that assertive *yum*, which is absent from Kalmyk, might exist in Qinghai Oirat (Deedmongol), i.e.

233   the south-easternmost dialect of Western Mongolian.

234   Examples from SC are transcribed phonemically with a transcription standard that draws

235   both from conventional transcription systems for Cyrillic and from IPA. The transcription of

236   examples from written sources is adjusted to better represent consonant phonemes and

237   syllabification, though the non-cliticization of *yum* in most writing is retained faithfully even

238   though it may contradict syllabification rules of the spoken language. The glossing and translation

239   of examples from other sources are always ours. In the translations of examples, parentheses

240   mark utterances present in the source that, in order to save space, are only given in translation,

241   or information that would be inferred from such passages. Brackets indicate information that was

242   not (literally) present in the source material or that resisted literal translation. Descriptions of the

243   overall context are given in curly brackets.

244

245   **3. Uses of *yum*: an overview**

246

247   In this section, we shall provide a brief overview of the synchronic network of *yum*-based

248   constructions and its historical implications. As a starting point, let us take a look at Table 2 that

249   shows the frequencies of *yum*-based constructions found in SC:

250

251   *[intended location of Table 2]*

252

253   Apart from its referential uses, *yum* occurs in two constructions in which it falls within the scope

254   of a sentence-final existential copula: action denial ('it is not the case that I V-ed', as discussed

255   below in (60)) and presumptive reasoning, as in (2e). Here, *yum* would appear to function as

256   matrix clause subject (i.e., if one would try to (mis)interpret *ter id-seŋ yim* in (2e) as 'the issue that

257   she ate') and simultaneously as the head of a preceding attributive clause (*yim bai-n* 'there is a

258   case / there are cases'). However, this analysis is incorrect since regular attributive clauses in

259   Central Mongolic allow for genitive subjects (i.e. ***tüün-ii** id-seŋ yim*), while this construction

260   requires a nominative subject (as in (2e)). Consequently, *yum* in these constructions is not a

261   subordinating element (anymore), but functions as an epistemological modifier of the matrix

262   clause in a syntactically flat structure. The two uses differ with regard to the morphological status

263   of *yum*: In the action denial construction, *yum* is a phonotactically distinct word (a particle), while

264   in the presumptive reasoning construction it may be cliticized to the preceding element. YUM as

265 a clitic or particle without a copula can also be found in the final position of subordinate clauses

266 (e.g. conditionals and similatives) or the matrix clause. These uses are mostly assertive /

267 presumptive / presuppositional and constitute the main research object of this paper. Among the

268 clause-final uses, there are also those in which *yum* and the subsequent element have turned into

269 a compound word. In such cases, the assertive meaning of *yum* is partially reinterpreted in terms

270 of a new clause-connecting function.

271 Since the main goal of this paper is a synchronic analysis, we have opted to discuss the

272 constructions at hand in an order that is semantically motivated: We first discuss assertive and

273 presumptive uses of *yum* in sentence-final position (§6), then its presumptive reinterpretation in

274 clause connection (§7), and finally the presumptive uses of copula-assisted *yum* in main clauses

275 (§8). Uses that don't fully fit into these classifications (action denial, 'or', 'because') will briefly be

276 mentioned in contexts where they structurally fit in.

277

278 **4. The morphological status of *yum***

279

280 Morphologically, as has been mentioned in §3, *yum* can morpho-syntactically take the form of an

281 independent noun, a particle that (in contrast to the noun) participates in syntactically flat

282 structures, and a clitic that is attached to the preceding element. Morpho-phonologically, this

283 development towards a bound morpheme is gradual and crucially depends on the frequencies of

284 individual morphological collocations. As a noun or particle, *yum* takes the form /jʊm/, i.e. [jʊm]

285 or [jim]. The clitic, by contrast, has the phonological form /jin/, but its realization differs somewhat:

286 When attaching to adjectives or nouns, it is [jin] or vowel-harmonic [jAn]. When attaching to the

287 participial suffixes [sǍŋ] (perfect/established past), [tǍg] (habitual) and [V̌x] (nonpast), /jin/ is

288 mostly realized as [in], and the codas of the perfect participle (obligatorily) and habitual participle

289   (optionally) drop, along with their non-phonemic reduced vowels (yielding [sin], [t(g)in] and [xin]).

290   Alternatively, speakers sometimes retain the full form [jin], but connect it to the preceding

291   participle with an epenthetic [i] (yielding [sijin], [tijin] and [xijin].[8] When attaching to the

292   established present form *baigaa* [pæː-ɢɑ] of the copula, *yum* only takes the form [n]. The final [n]

293   of the clitic is often realized as nazalization of the preceding vowel (e.g. [sĩ] instead of [sin]). If

294   retained, it is realized as [m] before bilabials or in intervocalic position.

295   While this phonetic variation can be described more or less objectively, the answer to the

296   question of how many morphemes should be postulated is an analytical decision. Different

297   Mongolian orthographies answer this question differently. In normative Cyrillic Mongolian as

298   used in the Mongolian state, reflexes of Common Mongolic *jɑʊmɑ are always rendered as юм

299   <yum>, though reduced forms occasionally make it into writing. In Contemporary Written

300   Mongolian as used in Inner Mongolia, a noun ᠶᠠᠭᠤᠮ᠎ᠠ <yaɢum-a> and a particle ᠶᠤᠮ <yum> are

301   distinguished. The approximate interdependence of cliticization and grammatical construction,

302   summarized in Table 2, provides evidence for two morphemes. This evidence is most

303   straightforward with participles which tend to require the use of the cliticized form =*iin* (unless in

304   very formal speech or in reading aloud), while nouns and adjectives (particularly those not derived

305   through the comitative suffix -*tAi*, cf. (7a)) more often co-occur with a de-accented, but

306   segmentally unreduced particle form.

307   In the remainder of this paper, we will use "YUM" as a shorthand for all non-referential uses

308   of *yum*/=*iin* irrespective of morphological form. In examples, we will disambiguate it by using the

---

[8] Judging from the transcription of SC, this longer variant might be most common in subordinate clause contexts, but since the transcription was not made to catch such quasi-allophonic variation, this might also be an artefact of the transcription process.

309   glosses "thing" and "ASS[ERTIVE]". In contexts that might arguably contain a non-referential noun,

310   we will occasionally use "NMLZ" (for "nominalizer") as an alternative glossing.

311

312   **5. Referential uses**

313

314   In Khalkha Mongolian, referential *yum* is used to refer to physical entities as in (3) (n_{SC}=94) or non-

315   material entities (including activities and processes in addition to concepts) (n_{SC}=280)[9] as in (4)

316   and (5).[10] In most contexts, *yum* is devoid of particular evaluative connotations, though its

317   semantic indeterminacy can be exploited for intentional indirectness in taboo contexts (e.g. *yum*

318   *uu-* thing drink- in the sense 'drink vodka' or *ɢazr-iiŋ yum* ground-GEN thing 'anthrax').[11] When referring

319   to humans, animals or organizations (n_{SC}=16/94)[12] as in (6), the non-agentive (since inanimate)

320   core meaning of *yum* is reinterpreted in terms of pity or contempt. *Yum* is the most common

321   abstract noun for things and processes, being far more frequent than its semantically most similar

322   competitor *züil* 'event, process, issue, thing' (n_{SC}=37).

323

324   (3) *bi   yeröösöö   ix   **yum** züü-(y)ii   ge-ž   bod-[dog=güi]* (SC)

325   1SG   at.all   many thing   put.on(jewelry)-CHRT COMP-CVB   think-PTCP.HAB=NEG

326   'I don't think at all that I want to put on a lot of <u>things</u>.'

327   (4) *yer-eed   oŋ-ii   exen-d   am'dral xecüüd-xed   neg büx=(e)l   **yum***

---

[9] This number includes 16 ambiguous cases.
[10] This use is first attested in the Late Western Middle Mongol (15th century?) source *Muqaddimat al-Adab* (edited by Saito 2008) where an abstract noun *ya:ma* is used with reference to inanimate physical entities and in some instances to non-physical non-eventive entities such as knowledge. In this function, it is also attested in corpora of Khorchin [Hán et al. 2012/2016] (0,48%, n=169/35000) and Kalmyk [Baranova 2007-2014] (0,45%, n=35/7743), with token frequencies that resemble Khalkha (0,62%, n_{SC}=374/60032). *Yum* cannot refer to time which in Khalkha is covered by *üy* 'period' and *cag* 'time'.
[11] https://mongoltoli.mn/dictionary/detail/137961, retrieved 2017-01-17
[12] That is, the overall **n**umber of tokens in which yum refers to physical entities is 94, and of these 16refer to humans, animals or organizations

328   90-APPROXIMAL year-GEN beginning-DAT life   become.difficult-CVB.when one all=FOC   thing

329   *xecüüd-seŋ*. (SC)

330   become.difficult-PST.EST

331   'In the early 1990s, when life became more difficult, all <u>things</u> became more difficult.'

332   (5) *ta   bol   öör=öö telwiz-ees öör   **yum** xii-x=güi   bai-n=uu?* (SC)

333   2SG.HON TOP   self=RPOSS TV-ABL   other   thing   do-PTCP.NPST=NEG AUX-PRS.IMM=Q

334   'So you yourself are not doing <u>anything</u> else except for TV?'

335   (6) *gexdee manai-d   yadruu   **yum** bai-x=güi=l=dee*. (SC)

336   but   1PL.EXCL.GEN-DAT somewhat.poor thing   be-PTCP.NPST=NEG=FOC=DP

337   'But there are no impoverished <u>individuals</u> around here [i.e. at this place of ours].'

338

339   There are three properties that referential *yum* shares with all other countable nouns: it is a

340   prosodically independent word which can receive accent; it can occur in clause-initial or clause-

341   medial position, and it can be morphologically marked (n_{SC}=122) for number, case and possession

342   as in (7a-c) and reduplicated for distributional and attitudinal readings as in (7d), where *yum~xum*

343   helps to tone down the importance that the speaker assigns to the referents. Morphologically

344   unmarked clause-final *yum* is potentially ambiguous between referential and assertive uses (see

345   §6.1). Structural ambiguity is higher in writing, since normative orthography ignores cliticization

346   and accent and always represents *yum* as an independent word.

347

348   (7) (a) *end min-ii **yum(ɴ)-uud** dutuu   bai-(ɢ)aa=daa*. (SC)

349   here 1SG-GEN thing-PL   incomplete AUX-PRS.EST=DP

350   '[We have just moved.] Here, many of my <u>things</u> are missing.'

351   (b) *texeeree nögöö   xüüxed=čin' nuu-x   **yum-tai** bol-ž   exel-deg*. (SC)

352   so   aforementioned child=STC   hide-PTCP.NPST thing-COM   become-CVB begin-PTCP.HAB

353   'So those children develop <u>things</u> that they [want to] hide.'

354   (c) *dur-tai*   **yum=aa**   *üz-(e)n* (SC)

355   like(desire-COM) thing=RPOSS   look-POT

356   'They watch the <u>things</u> they like'

357   (d) *tend-ees* **yum~xum**   *aw-aa=l* (SC)

358   there-ABL   thing~DISTRIBUTIVE take-CVB.PFV=FOC

359   'From there, we bought <u>this and that</u> and ...'

360

361 In most cases (n_{SC}=229/374, 61.2%), *yum* hosts short attributes consisting of one or two elements

362 such as quantifiers in (3) and (4), determiners in (5), adjectives in (6), comitative nouns in (7c),

363 and participles in (7b). Since words from these word classes (with the exception of some

364 determiners) usually cannot refer to participants purely on their own, using *yum* as the most

365 abstract noun available is an efficient way to create participant-referring noun phrases for non-

366 human(-resembling) patients and themes. *Yum* is not as commonly used for agents and humans

367 (which can be referred to by specific nouns or by the general noun *xüŋ* 'person, someone'), places,

368 times (for which the abstract nouns *üy* and *cag* exist along with converbal expressions) and events,

369 which are more commonly nominalized by participles followed by either case suffixes as in (8a)

370 or personal possessive clitics as in (8b). On the other hand, the number of cases in which *yum* is

371 preceded by more complex attributes (n_{SC}=33) is relatively small.

372

373   (8) (a) **En xereg xezee ɢar-sŋ-iig**   *med-(e)x=güi*. (IC)

374   this event when   occur-PTCP.PRF-ACC know-PTCP.NPST=NEG

375   '[I] don't know <u>when this event occurred</u>.'

376   (b) **Tiim xereg ɢar-saŋ=n'**   *üneŋ*. (IC)

377   such   event   occur-PTCP.PRF=3POSS true

378   'It is true <u>that such an event occurred</u>.'

379

380   Apart from this, *yum* can also occur without attributes (n_{SC}=64). In the case of direct objects, this

381   results in quantification since the presence/absence of the accusative case indicates

382   specificity/non-specificity (see Guntsetseg 2016), as (9) illustrates.

383

384   (9) *xüŋ-tei*   **yum** *yar'-ž*   *čad-ax=güi*. (SC)

385   person-COM thing   speak-CVB can-PTCP.NPST=NEG

386   'I cannot talk with people [<u>even a bit</u>].'

387

388   Guntsetseg (p.c., 2017-04-03) further suggests that *yum* might occasionally be used as a dummy

389   argument in thetic (or "unanchored") sentences to fill an obligatorily overt argument position.

390

391   **6. Sentence-final YUM**

392

393   Apart from its pimary function as an abstract referential noun, Khalkha YUM can be used in clause-

394   and sentence-final predicates in a modal or interactional function.[13] It is highly frequent in such

395   functions, being present in a total of 3-15% of sentences in the theatre plays of Möŋx-Amɢalaŋ

396   (1996)'s corpus and constituting or being part of 2,1% of words in SC (n=1282). This section will

397   focus on the most frequent among these secondary uses, which is to either reinforce

398   assertiveness in statements or presume certain states-of-affairs in questions. In SC, this usage is

---

[13] *Yum* did not develop non-referential uses in Kalmyk and Khorchin at all. In Khorchin, though, it converted to a verb
jim- 'do so' (Bayancoytu 2002: 277-278).

399   attested in 777 of 4248 finitely used positive indicative verbal predicates (18%) and in overall 145

400   nominal predicates. Morpho-syntactically, sentence-final YUM forms a grammatical class unto

401   itself in that it occupies a dedicated slot in-between nominals (participles, nouns, adjectives,

402   existential-identificational particles) and all other sentence-final (modal or discourse) particles. It

403   patterns with sentence-final particles in that it cannot be used on its own as an answer to a

404   question (Umetani 2013: 305, 308-310). Since most modal particles (which express epistemic and

405   evidential notions without referring to the relationship between the interlocutors) are fossilized

406   copulas, it is indeed expected that they should follow YUM as an erstwhile noun under SOV word

407   order. Due to its nominal provenience, it is also expected that YUM cannot follow finite (indicative

408   or mood) suffixes. Its position within the overall system of sentence-final particles, with slight

409   simplifications, is shown in Table 3 (building upon Kang 2003: 41):

410

411   *[intended location of Table 3]*

412

413   Table 3 shows that YUM combines with participles which express established information, while

414   it does not combine with indicative declarative suffixes that express immediate realization (see

415   Brosig & Skribnik 2018) or with non-indicative mood suffixes (such as imperatives, hortatives, or

416   preventives). It may precede most modal particles with their various epistemological meanings

417   (e.g. *baix* 'possibly, probably', *šiw* 'apparently', *bilee* [recollection], *až* [indirect evidence]), which

418   are in turn followed by epistemic discourse particles that intersubjectively allocate epistemic

419   authority to the speaker (*=šüü*) or addressee (question particle *=UU*). [14] Its meaning must

---

[14] The diachronicallly heterogeneous discourse particle *=AA* fulfills a variety of stance-related functions such as expressing politeness (Brosig 2015: 61,75-76), enthusiasm (Kullmann&Tserenpil 1996: 337) or insistence, but more research is needed for a somewhat consistent account of its functions.

420   therefore be compatible with both epistemic and intersubjective notions, though there might

421   certainly be some conventionalized interaction.[15] In the following, our discussion will only take up

422   the distinction between unmarked declaratives and interrogatives marked by =*uu* and =*be*, while

423   leaving aside the synergies that YUM develops with the other sentence-final particles or even with

424   particular participles.[16]

425

426   6.1 Sentence-final YUM in declaratives

427

428   Declarative sentence-final YUM ($n_{SC:PTCP}$=465, $n_{SC:OTHER.NOMINALS}$=36) indicates the **speaker's full**

429   **subjective commitment** to an utterance, thus increasing its assertive force. Mönx-Amɢalaŋ

430   (1996), in his detailed but decontextualized analysis, describes this in terms of individual meanings

431   such as *xüč nemeŋ batalsaŋ / oncloŋ batalsaŋ / čuxalčlaŋ batalsaŋ* 'assert while adding force /

432   particularly singling out / emphasizing as important', *medeeleŋ sonsɢosoŋ* 'inform and familiarize'

433   and *ünemšüüleŋ itgüülex* '(try to) convince', which can be interpreted in terms of the different

434   conversational purposes that motivate such commitment. In contrast to a discourse particle

435   combination such as *š*=*d* (i.e. *šüü*=*dee*) which indicates the speaker's and a third party's higher

436   epistemic authority and thus "objective truth", the commitment conveyed by YUM appears to be

437   the speaker's own evaluation or subjective truth, and YUM indeed often occurs in contexts where

438   the speaker talks about her own feelings or intentions such as (10) or tries to explain and justify

439   her deeds as in (11). Informants suggest that the speaker actually expresses her feelings by

440   uttering (10), while she would merely report them if YUM was absent. A similar situation can be

---

[15] For instance, *šiw* combines with -*x*/YUM, hinting at a distributional pattern similar to the one discussed in §7.2, while *dag* tends to combine with -*s*=*iin* or *bai-n*.

[16] The interpretation of *yum* in combination with participles, the modal particles *dAg*, *bil(ee)* and *sAŋ*, and all discourse particles except *xöö* is, however, discussed by Mönx-Amɢalaŋ (1996).

441 observed in (12) where the absence of YUM would either result in a mere statement about the

442 past (which would not be felicitous in the context of the teachers' meeting during which this

443 sentence was uttered) or about an entire lack of concern on the part of the speaker. In the pattern

444 *-x ge-*PAST which denotes subject intention (see especially Song 2002), the YUM-based variant

445 requires a first-person subject (Mööŋx-Amɢalaŋ 1996: 29) as in (13).

446

447 (10) *en čemdaŋ-g odoo bi ix* **xairal-d=iin**. (SC)

448 this suitcase-ACC now 1SG very love-PTCP.HAB=ASS

449 '[The things that I am most proud of in my home, the things that I am most happy about

450 and would like to show people are my books. When I came from London, I brought seven

451 suitcases of books.] This suitcase now I <u>love</u> a lot. [In this suitcase, I used to put my most

452 important books.]'

453 (11) Ranma: Hey...Where's Ryoga?

454 Genma: Oh, he took the Kinjakan [a kind of magical staff that can be used

455   like a supersonic motorcycle] to get some food.

456 Ranma (dazzled): Eh...Ryoga...

457 Muus (dazzled):   Took off with the Kinjakan...

458 Ranma (angry):   And you just let him go?

459 Genma (sweating): Um... I wanna eat sometime...

460   *Mm... bi yum* **id-meer** **bai-s-iin**...[17]

461   um   1SG thing eat-wish.CVB AUX-EST.PST=ASS

462   (Lit. Um ... I wanted to eat something.)

---

[17] Takahashi, Rumiko. 1996. Ranma½ 37(8): 123. Japanese-English translation of this manga by Jason Bridgmon, English-Mongolian translation by Atagu (2013).

463   (12)*Bi zügeer aŋɢ _daa-saŋ_bagš-iiŋ xuw'd tan-d   xel-ye ge-ž bod-soŋ   yum*. (Mahal 2001: 80)

464   1SG   just   class_teacher   for   2SG.HON-DAT say-HORT QV-CVB think-PTCP.PRF ASS

465   'I just thought, as the class teacher, that I should tell you this.'

466   (13)*Noyo<o>ŋ, bi ta(n)-d   belegl-ex   **ge-seŋ**   yum*. (Möŋx-Amɢalaŋ 1996: 29)

467   lord<EMPH>   1SG   2SG.HON-DAT   present-PTCP.NPST   say-PTCP.PRF   ASS

468   'Lord, I want to offer a present to you.'

469

470 However, YUM is not limited to such contexts, and it is perfectly possible for a speaker to use YUM

471 about events in which the speaker is not a participant, such as *soxtuu yaw-x=iin*   '[they] would

472 always be drunk' in (14), or also when adding emphasis (i.e. conveying the writer's conviction) in

473 newspaper texts. In emotionally loaded contexts such as (14), where a speaker tries to convince

474 his addressee of an opinion that contradicts the published opinion, YUM may be used in several

475 consecutive clauses, while such sustained use is precluded in formal style. Interestingly,

476 informants interpret *taar-x=iin* in (14) in two ways, namely either as the speaker's emotional

477 conviction that he wants immediately enforced or as the speaker's subjective opinion that he

478 seeks support for but which the addressee need not (yet) share.

479

480   (14)*odoo en süüliiŋ … šuud xel-xed   süüliiŋ dolooŋ naimaŋ žil   saixaŋ ox'-duud*

481   now   this recent   direct say-CVB.when last   seven   eight   year beautiful girl-PL

482   *dandaa soxtuu **soxtuu yaw-x=iin**.   en baar zaar-uud-iig bügd-ii=n xaa-wal*

483   always   drunk   drunk   go-PTCP.NPST=ASS   this baar   RED-PL-ACC   all-ACC=3POSS close-CVB.COND

484   ***taar-x=iin**.   bi šuud **xel-žai-(ɢ)aa=m**   šüü. en xüüxd-iiŋ-x=aa   tölöö.*

485   fit-PTCP.NPST=ASS 1SG direct say-PROG-PTCP.PRS=ASS   DP   this child-GEN-NMLZ=RPOSS for

486   *ard_tümn-ii xüüxed dund surɢuul' oydl-iiŋ   surɢuul' yaaɢaad **tögs-č***

487   people-GEN   child   middle school   sewing-GEN school   why   finish-CVB

488   ***bol-d=güi=n.***   *neg xüüxed yaaɢaad ɢutl-iiŋ surɢuul' tögs-ööd ɢutal*

489   may-PTCP.HAB=NEG=ASS one   child   why   shoe-GEN school   finish-CVB.PFV shoe

490   ***xii-ž*** *yaaɢaad* ***bol-d=güi=n***. (SC)

491   make-CVB why   may-PTCP.HAB=NEG=ASS

492   'Now these last … saying it outright, beautiful girls <u>would</u> always <u>be drunk throughout</u>

493   these last seven, eight years. One <u>ought to</u> close down these bars, all of them. I'<u>m</u>

494   <u>speaking</u> straight. For the sake of these children! Why <u>may</u> the children of our people <u>not</u>

495   <u>graduate</u> from middle school or sewing school? Why <u>may</u> some child <u>not graduate</u> from

496   cobbler's school and make shoes?'

497

498   One common type of usage that Mukai (2001: 85-86) mentions but leaves unclassified is in

499   expository sentences like (15). In our account, these are explained in terms of the writer's

500   subjective commitment. When explaining such sentences, informants emphasize the speaker's or

501   writer's expertise or subjective certainty. Due to this subjective commitment, the proposition of

502   the sentence is understood as soundly asserted and not in need of further explanation, something

503   that occasionally leads to an interpretation as a sentence that in itself constitutes an explanation.

504   We thus interpret Jīngāng's (2007) explanation in terms of 'reason why something occurred' as a

505   relatively frequent connotation, while our analysis leaves no room for his interpretation of YUM

506   as marking 'generally known facts'.[18,19]

507

---

[18] Subjective examples like (10)-(13) obviously don't convey generally known facts, and neither do expository examples like (15) when interpreted by informants. In (14), the presence or absence of YUM mostly correlates with the personal confidence and conviction of a speaker who uses YUM in an attempt to convince the audience.

[19] The reader may notice that we do not cite any examples from Jīngāng (2007). The reason for this is the overall sparsity of examples in Jīngāng's paper with regard to this part of his semantic analysis (arguably only his examples (30)-(34)), the lack of contextualization of these examples, and structural features in his examples (31)-(33) that were rejected by some of our informants during elicitation (presumably for dialect-related reasons).

508   (15) *Šizofreni   šinž_temdeg oloŋ   yanz-iiŋ **ilrel-tei**   **yum**.* (Mukai 2001: 85)

509   schizophrenia sign   many   kind-GEN   manifestation-COM ASS

510   'Schizophrenia has many different manifestations.'

511

512   The specific use of YUM as a 'focus adjustment marker' claimed by Mukai (2001) is illustrated in

513   (16). Here, the presence of YUM is claimed to shift the focus from the main predicate to the

514   subject, *bi* 'I', as contrasted to some other agent. While informants tended to mention the

515   speaker's sincerity and the explanatory power of the sentence when discussing the contribution

516   of YUM to this sentence, they indeed were unwilling to accept any interpretation that puts a focus

517   on the predicate. However, this is also due to the predicate itself, since the Mongolian verb *zur-*

518   as a general term also covers drawing and other techniques, so that virtually any normal picture

519   must have been produced by an activity that can be described by this verb. However, if the verb

520   is changed to *xuul-* 'copy > plagiarize', all informants immediately assume a contrast with a

521   predicate, namely *bič-* 'write [oneself]'. Since this operation does not change the overall structure

522   of the sentence, it seems that the contribution of YUM cannot be explained purely in terms of

523   focus even in this type of sentence.

524

525   (16) *Ünen-iig xel-exed   BI zur-saŋ   yum*. (Mukai 2001: 81)

526   truth-ACC   say-CVB.when 1SG paint-PTCP.PRF ASS

527   'To say the truth, "I" painted it [myself].'

528

529 Among non-participial predicates combining with YUM, adjectives[20] as in (17) are prevalent. In

530 such predications, the status of YUM as particle or noun in written contexts is potentially

531 ambiguous. As seen in (18), occasional co-occurrences with nouns, where *yum* is unambiguously

532 interpreted as a sentence-final particle, are attested as well.

533

534   (17)*za   za   **sürxii yum=aa***. (SC)

535   well well terrific thing/ASS=HON

536   1. 'Well well, that's <u>terrific</u>!'

537   2. 'Well well, that's <u>a terrific thing</u>.'

538   (18)*en   büxeŋ kaimr-iiŋ **ür_düŋ yum=aa***. (SC)

539   this all   camera-GEN result   ASS=HON

540   'All of this [i.e our increased ability to detect crime] <u>is the result</u> of the cameras!'

541

542 6.2 Sentence-final YUM in questions

543

544 YUM in questions differs from its declarative counterpart both in terms of function and frequency.

545 In functional terms, interrogative YUM is epistemically weaker than declarative YUM: While the

546 speaker uses declarative YUM to designate a given statement as her own personal conviction and

547 thus conveys (subjective) epistemic certainty, in interrogatives this conviction may get

548 reinterpreted in weaker terms such as expectation of certain facts or the anticipation of a certain

549 assertion on the part of the addressee. In other words, the speaker may use YUM to indicate that

550 she, at the time of speaking, **presumes** a certain state-of-affairs to hold (or that she expects the

---

[20] In Mongolian, the comitative case *-tai* can still express its original possessive meaning, e.g. *son'orxol-toi* 'having interest > interesting, exciting' or *xereg-tei* 'having necessity > must'. The words thus derived have an adjective-like distribution and, for the purpose of this paper, are treated as adjectives.

551 authoritative addressee to be committed to a certain state-of-affairs). Consequently, YUM can be

552 used in **interrogative** structures that are not meant to improve the speaker's understanding of

553 the world with regard to the content that the question, if understood as such, would elicit. Apart

554 from exam questions, this is also true of **exclamatives** and **rhetorical questions**, where a certain

555 state of affairs is not only presumed, but fully presupposed. Alternatively, the speaker may take

556 a disinterested stance with regard to the information asked about since it is already presumed.

557 In terms of frequency, interrogative YUM in SC accounts for 41% ($n_{SC:PTCP}$=321/786) of the

558 overall sentence-final tokens of YUM with participial predicates and for 71% ($n_{SC:NOM.PRED}$=89/125)

559 of sentence-final tokens with other nominal predicates. While the ratio of interrogative to

560 declarative sentences in SC cannot be calculated based on its current annotation, it is safe to

561 assume that is must be below 1:2. This means that the relative frequency (and prominence) of

562 YUM in interrogatives is much higher than in declaratives. The diverging relative frequencies with

563 participles and other nominals also indicate that the word class of the predicate's core lexeme

564 plays a certain role.

565 As stated, one of the two main functions of YUM in questions is expressing that the speaker

566 anticipates a certain answer (*lawlaŋ batlaž asuux* 'asking to consult and confirm', Möŋx-Amɢalaŋ

567 1996: 36). The presence of YUM usually distinguishes such questions from open questions (cf.

568 Kang 2003: 31) where such anticipation is lacking. In (19)(a), for instance, the speaker uses YUM

569 when asking about a state-of-affairs that he had already guessed on the basis of his observations

570 while being a guest in the home of the family he is visiting. In contrast, a visitor with no previous

571 knowledge about the addressee's family who thus wouldn't anticipate any particular answer,

572 would ask the same question without YUM, as in (19)(b) from the Frequently Asked Questions

573 section of a personal website. The use of YUM in polar questions can entail different nuances, for

574 instance offence (Möŋx-Amɢalaŋ 1996: 36) as in (20), or confirmation of newly learnt information

575 along with a shade of doubt (Mönx-Amɢalaŋ 1996: 37) as in (21). In the latter case, a pre-existing

576 personal conviction is replaced by what the speaker assumes she is entitled to believe given the

577 conversational context. In (22), informants hold that Speaker B had not anticipated to be given a

578 present earlier on but that, at the time of speaking, she does anticipate to receive this present

579 and, thus, a positive answer. (23) illustrates an alternative question in which the speaker suspects

580 that the addressee has actually broken the object, thus anticipating a negative answer to the first

581 and a positive answer to the second alternative.

582

583 (19)(a) *ta   ail-iiŋ   **tom yum=uu**, tee*? (SC)

584   2SG.HON family-GEN large ASS=Q   INTERJ

585   'You are the oldest (child) of the family, right?'

586   (b) *Ail-iiŋ   **tom=uu**?*[21]

587   family-GEN large=Q

588   'Are you the oldest [child] of the family?'

589 (20)*Či   nadad xair=güi **bol-čix-soŋ**   yum=uu?*[22]

590   2SG   1SG.DAT love=NEG become-COMPL-PTCP.PRF ASS=Q

591   'Have you lost your love for me?'

592 (21)*Aa   öndör Ceweeŋ-ii **xüü** yum=uu?* (Mönx-Amɢalaŋ 1996: 37)

593   INTERJ large   NAME-GEN   son ASS=Q

594   'Oh, he's the son of Big Ceveen?'

595 (22)A: *en   moŋɢol-oos aw-saŋ   beleg.*

596   this Mongolia-ABL   take-PTCP.PRF present

---

[21] https://ask.fm/anukaorji, ca. July 2016, retrieved 2017-03-13. From a question-answer section.
[22] Mönxsaihan Xoŋgorzul (reporter): Duučin B.Amarxüü, Aziin Top model' O.Ar'uunzul: Am'draliiŋ utaɢ učir xüüxed yum bain. http://www.ubs.mn/news/12275, 2015 or 2016, retrieved 2017-03-13

597   'This is a present [that I] brought from Mongolia.'

598   B: *nadaa ög-č   bai-(ɢ)aa yum=uu?*   (Mukai 2001: 78)

599   1SG.DAT   give-CVB AUX.PRS.EST ASS=Q

600   'Are you giving [it] to ME?'

601   (23)*Xaɢar-saŋ   yum=uu, xaɢal-saŋ   yum=uu?* (Mukai 2001: 78)

602   break(INTR)-PTCP.PRF ASS=Q   break(TR)-PTCP.PRF ASS=Q

603   'Did it break, or did you break it?'

604

605 When asking YUM-based polar questions, the speaker thus presumes that the state-of-affairs

606 asked about either holds or does not hold. In content questions, no possible state-of-affairs that

607 might serve as an answer is put into words, but the speaker still anticipates a specific answer or

608 type of answer. For instance, in (24), the speaker is asking about the age of a small child which

609 she can approximately guess. In (25), a police officer asks a suspected thief about a phone that he

610 presumes the suspect to have stolen from one of the (not specifically known) customers of a large

611 street market.

612

613   (24)*xüüxed=čin' odoo xedeŋ   **sar-tai   yum=be**?* (SC)

614   child=STC   now   how.many month-COM ASS=Q

615   'How many months is your child now?'

616   (25)*en xeŋ-ii **utas yum=be**?* (SC)

617   this   who-GEN phone ASS=Q

618   'Whose phone might this be?'

619

620   When the speaker considers her own knowledge as authoritative, YUM can be used in "insincere"

621   questions such as topical questions meant to draw the addressee's attention that the asker later

622   answers herself as in (26) or exam questions like (27). In this type of questions, the addressee is

623   not necessarily expected to know the full answer, but is at least presumed to hold relative

624   knowledge that could be used when approximating the actual state-of-affairs.

625

626   (26) *Yaaɢaad moŋɢol   aaw   öglöö   cai=(ɢ)aa uu-xd=aa   ger-iiŋ*

627   why   Mongolian father morning tea=RPOSS   drink-CVB.when=RPOSS yurt-GEN

628   *baruuŋ xoimor-t   suu-dag   yum=be?* (IC)

629   western   seat.of.honour-DAT sit-PTCP.HAB ASS=Q

630   'Why does a Mongolian father, when he drinks his tea in the morning, <u>sit</u> at the western

631   seat of honor {at the north side} of the yurt? (As science has confirmed, [the reason is]

632   that the beautiful, auspicious morning sun is first reflected on the north-western wall of

633   the tent. (…))'

634   (27) *Üün-d=čin' yamar dörwöŋ nügel bagt-dag   yum=be?* (IC)

635   this-DAT=STC   what   four   sin   belong-PTCP.HAB ASS=Q

636   'What four sins belong here (i.e. to the sins of the tongue)?'

637

638   The presumptive meaning expressed by YUM makes it a viable device in fully presuppositional

639   question types such as rhetorical questions and exclamatives. In the absence of a typologically

640   viable definition, we will consider as a "rhetorical question" an utterance that is structurally coded

641   as a question, but is primarily meant to implicate a certain state-of-affairs. The rhetorical question

642   retains a hybrid status as a quasi-statement, short of a direct assertion, and potentially with the

643   option of being interpreted as sincere. That it is not intended as a sincere question is often only

644 made transparent through formal devices such as intonation pattern, special particles or verbal

645 (subjunctive) mood. Functionally, rhetorical questions in several Romance and Germanic

646 languages often communicate the speaker's guarded reserve towards the truth of the proposition

647 or conversely the status of the implied proposition as intersubjectively self-evident, along with

648 more specific propositional attitudes (like reproach, indignation, protest, wonder, perplexity,

649 dismay or emphasis) (Schmidt-Radefeldt 1977). Since the speaker asking a rhetorical question

650 implicates that she presupposes a certain state-of-affairs, one would expect rhetorical questions

651 in Khalkha to often feature YUM. For rhetorical content questions, this prediction generally seems

652 to hold. In (28), YUM helps to prevent the question from being interpreted as sincere and to

653 implicate a certain disregard.[23] YUM is also occasionally found in rhetorical polar questions such

654 as (29). In this particular case, the variant without YUM, lacking the courtesy of implying a

655 (negative) answer, would be perceived as a real question and thus as offensive. However, polar

656 rhetorical questions are formed from a wide range of source patterns including *ge-ž=üü* 'Are you

657 saying (as I infer) that …?', *gež bod-(o)ž bai-n=uu* 'Are you [really] thinking that …?' and *biš=üü*

658 'Isn't it rather the case that …' (cf. Gāo 2013) which mostly build on an incredulous stance towards

659 the addressee's thoughts rather than a presupposed common ground between the interlocutors.

660

661 (28) *1000 xüŋ   xičeelel-deg   sport-oor awraɢ   bol-(o)xod*

662   1000   person practice-PTCP.HAB sport-INS   champion become-CVB.when

663   *yuu=n'   **tiim sürtei bai-dag   yum=be**. (IC)*

664   what=3POSS so   terrific   AUX-PTCP.HAB ASS=Q

---

[23] Möŋx-Amɢalaŋ (1996: 34-35) analyzes such constructs in modal rather than illocutionary terms, but this probably exaggerates the extent to which these rhetorical patterns are grammaticalized.

665   'What <u>is so great</u> if somebody becomes champion in a sport that is practiced by [merely]

666   1000 people?

667   → There is nothing special about becoming the champion of such a sport.

668   (29)*či*   **teneg yum=uu**? (Gāo 2013: 16; IC)

669   2SG   stupid ASS=Q

670   'Are you stupid?'

671   → While I know that you are not stupid, what you are doing would suggest you are, so

672   stop it!

673

674 Exclamatives express the speaker's affective stance and surprise towards the extent of some

675 scalar property. While the specific extent of the scalar property is not precisely clarified, its

676 applicability as such is presupposed (cf. Michaelis 2001). In Khalkha, the most common pattern to

677 form exclamatives is through content questions containing scalar question words, especially *yaa-*

678 *saŋ* do.what-PTCP.PRF 'how' / 'has done what'. Whether YUM is present in exclamatives largely

679 depends on the word class of the predicate.[24] YUM is usually required in questions based on bare

680 adjectival predicates (n_{IC}=122/135, 90%) as in (30)[25] and participles (n_{IC}=81/110, 74%) as in (31),

681 though these also arguably also allow for alternative strategies of structural subordination

---

[24] The following conclusions are drawn from an analysis of collocations of *yaasaŋ* followed by one of the adjectives *aixtar ~ aimaar* 'awful, notable', *aztai* 'lucky', *ert* 'early', *ewgüi* 'awkward', *ix* 'much', *muuxai* 'ugly, bad', *oloŋ* 'many', *saiŋ* 'good', *saixaŋ ~ goy* 'beautiful, nice', *son'oŋ* 'interesting', *teneg* 'stupid', *udaaŋ* 'slow', *üneŋ* 'true', *üntei* 'expensive', *xatuu* 'hard', *xecüü* 'difficult' and *xöörxöŋ* 'cute' (n_{IC}=439).

[25] The adjective-based examples lacking YUM show rather particular contexts, e.g. in two cases very long preceding subject clauses. A simple constructed YUM-less exclamative sentence, *En yaasaŋ xačiŋ we?* (presumably intended) 'How strange this is!', is considered ungrammatical by Kang (2003: 32) and (if we discount a marginally possible use as a self-directed sincere question) considered unacceptable by our informants, as would be (30) without YUM.

682   (n_{IC}=27/110, 25%).[26],[27] If YUM is present, the content question clitic =*be* may be present or absent.

683   With nouns as in (32), YUM is almost never found (n_{IC}=1/141), but =*be* is used consistently

684   (n_{IC}=137/141). That is, the structural conditions that obtain here differ from those of simple YUM-

685   based assertives. Exclamatives also allow for finite verb forms (n_{IC}=44), mostly the surprised past

686   -*w(=aa)* and an emphatic form of the potential, -*n=aa*, in which case YUM is absent as well (which

687   is expected since even most older sentene-final particles that developed from nouns are

688   incompatible with these finite suffixes). In conclusion, it appears that forms that lend themselves

689   to nominalization like adjectives and participles by and large require YUM in exclamatives, while

690   nouns don't require these and the requirement is waived for finite verbs (which remain in place

691   and do not e.g. have to be replaced by participles, as would be the case under negation).

692

693   (30)*Ter xödöö(n)-ii   zurag yaasaŋ **saixan yum**=bee* ... (IC)

694   that countryside-GEN picture how   beautiful ASS=Q

695   'How beautiful these pictures from the countryside are!'

696   (31)*Ta   yaasaŋ ɢoy **žüžigle-seŋ   yum**=be*? (IC)

697   2SG.HON how   nice   play.role-PTCP.PRF ASS=Q

698   'How beautifully you played your role [in the theatre play/movie]!'

699   (32)*Yoooo … Yaasaŋ aimar zurag=we*? (IC)

700   INTERJ   how   awful   picture=Q

701   'Eww/Eek … What a disgusting picture!'

---

[26] These alternative patterns are based on either the quotative verb *ge-* (especially *geeč* 'say!') or the conditional particle *bol* 'if > I am wondering', though it is not in all cases clear whether these should be classified as exclamatives or rather as bona fide miratives (*geeč*) and actual questions for the extent of a feeling (*bol*). Examples are *Ter üy-d yaasaŋ ix gancaard-saŋ ge-eč* that time-DAT how much feel.lonely-EST.PST QI-IMP(imploring) '(Imagine) how much [I] was feeling lonely in this time!' and *erx.čölöö-tei bol-ɢo-wol yaasaŋ ix bayarl-ax bol=doo* freedom-COM become-CAUS-COND.CVB how much feel.happy-FUT.PTCP MP=DP 'If you would give them freedom, how happy they would feel!'

[27] This count excludes the resultative participle -*aa* as attached to stems other than *bai-*. In our materials, -*aa* only combines with =*be* and never with YUM or other formal subordinators (n_{IC}=25).

702

703 In some contexts, YUM in regular (i.e. non-rhetorical) questions is used not so much to signal

704 presumption, but DISINTERESTEDNESS, in the sense of a lack of specific gain that the speaker could

705 derive from the answer. This use may have originated from presumptive questions where the

706 speaker is only expecting a confirmation rather than any real gain in world knowledge from the

707 answer. Weak epistemic questions with YUM=*uu* like (33), which convey guesses that often

708 implicate that the validity of the guess is of no concern to the speaker, would thus constitute an

709 intermediate stage between presumptive and disintested questions. An instance of a

710 disinterested question would be , asked by a reporter in the course of an at-home interview.

711 Judging from the wider conversational context, the reporter did not have any reason to expect a

712 positive answer to this question, but the reporter was mostly using YUM to signal empathy and

713 curiosity. In other words, "the speaker is really interested in the coming answer" (Street 1963:

714 160). In (35), disinterested YUM deemphasizes the negative consequences that the addressee's

715 behavior might have had on the speaker. Namely, it ensures that the sentence is not interpreted

716 as an insistence that the addressee should have taught the speaker frankly (Möŋx-Amᴳalaŋ 1996:

717 35). The expression of disinterest need not necessarily smoothen conversation, since the

718 presence of YUM can also be interpreted in terms of the speaker asking in a *xol xöndii baidlaar*

719 'distant, disattached/aloof way' (Möŋx-Amᴳalaŋ 1996: 32), i.e. as lack of both interest and

720 curiosity. In this type of usage, the original implication that the speaker already anticipates a

721 certain answer is lost, and disinterestedness becomes the main semantic contribution of YUM.

722

723 (33)A: *izrail-d xedeŋ xonog bol-ox=we*?

724 Israel-DAT how.many overnight.stay become-PTCP.NPST=Q

725 B: *oŋᴳc-oor bol boᴳ'on, end-ees boᴳoŋ yaw-ax=güi. end-ees=čin' oŋᴳc-oor*

726   plane-INS   TOP short   here-ABL   wagon   go-PTCP.NPST=NEG here-ABL=STC   plane-INS

727   *moskwo damži-ŋ   xor'oŋ xedeŋ   cag **yum=uu***. (SC)

728   Moscow   pass.through-CVB twenty   how.many   hour   ASS=Q

729   A: 'How many days does it take to Israel?'

730   B: 'It's not far by plane, there's no train connection from here. From here with a plane via

731   Moscow, it <u>would perhaps take</u> twenty odd hours?'

732   (34)*či en ger-iig öör=öö **xii-s=iim=uu***? (SC)

733   2SG   this house-ACC self=RPOSS do-PTCP.PRF=ASS=Q

734   'Have you built this house yourself?'

735   (35)*Tegeed bid(n)-iig öglöö oč-(o)xod   šuluu=xaŋ   zaa-(ɢ)aad ög-öx=čin'*

736   then   1PL-ACC   morning come-CVB.when plain=DIMINUTIVE show-CVB.PFV give-PTCP.NPST=STC

737   *yaa-saŋ   **yum=be***. (Möŋx-Amɢalaŋ 1996: 34-35)

738   do.what-PTCP.PRF   ASS=Q

739   'Then when we arrived there in the morning, <u>would it have been possible</u> to teach us

740   frankly?'

741   without YUM: '… <u>how would it have been</u> if you had [just] taught us frankly?'

742

743   Since YUM in polar questions not only communicates that the speaker presumes a certain state-

744   of-affairs to hold, but also allows the addressee to (more or less) unequivocally identify that state-

745   of-affairs, polar questions containing YUM, by extension, can also be used to suggest a

746   presupposed state-of-affairs as a possibility for consideration. Put differently, the question is not

747   addressed to a real interlocutor, but to an imagined ideal evaluator (cf. Tantucci 2017) who could

748   assess the reasonability of the claim. This usage is illustrated by (33) and by (36), the latter a self-

749   directed question from a monologue conveying epistemic possibility. In a cline of epistemic

750   markers, this use of YUM=*uu* would represent explicit non-commitment to a proposition,[28] while

751   *baix* in (37) would express an inclination to accept the proposition, and *biz* in (38) would signal

752   strong (intersubjectivized) commitment. In other words, epistemic probability increases from

753   YUM=*uu* via *baix* to *biz*.

754

755   (36)… *en* **xair yum=uu** *bi med-(e)x=güi yum=aa*[29]

756   this love   ASS=Q   1SG know-PTCP.NPST=NEG   ASS=DP

757   'Is this love? I really don't know.'

758   (37){Enumerating a number of behavioral traits of a life partner such as taking one's partner's

759   interests as being as important as one's own, not imposing demands in return, etc., the

760   writer concludes:}

761   *En=(e)l žiŋxen* **xair baix**.[30]

762   this=FOC true   love   MP

763   'Probably only this is true love.'

764   (38)*Bi=č či(n)-ii aŋx(n)ii* **xair biz=dee**[31]

765   1SG=FOC 2SG-GEN   first   love   MP=DP

766   '(You are my first love,) and isn't it so that I am your first love, too? (If not, say so!)'

767

768   The use of question-final YUM in contexts of disinterest has led to a certain epistemic ambiguity

769   or underdeterminancy of YUM. To this, we can add an additional observation. If content questions

770   contain YUM in our spoken data, it is very likely to occur in absolute sentence-final position (i.e.

---

[28] Möŋx-Amɢalaŋ (1996: 37-38) associates such constructs with the expression of either guessing or doubt, while we hold that it always denotes both notions, though the larger context may of course give prominence to one of them.
[29] http://www.cekc.mn/2013/10/хайр-ийм-хүчтэй-юм-уу, 2013-10-26, retrieved 2017-04-02
[30] http://www.ugluu.mn/190246.html, 2016-03-30, retrieved 2017-05-12
[31] http://www.cekc.mn/2014/02/Жаргахын-тулд-эсвэл-тэсэхийн-тулд-юм-болов-уу?, 2014-02-27, retrieved 2017-05-12

771 without being followed by any modal or discourse particle). This absolute sentence-final use

772 accounts for 73% (n_{PTCP}=135/185) of participial and 98% ($n_{NOMINALS}$=39/40) of other nominal

773 predicates in SC, while participle-based content questions without YUM generally require the

774 presence of a question particle.[32] This contrasts with the written data from IC, where bare YUM

775 is only found in 31% of content questions (n_{PTCP}=493/1596, $n_{NOMINALS}$=80/250), while the remainder

776 is followed by the content question particle =*be* or similar devices such as *bol* 'if > maybe?'[33] If the

777 differences between SC and IC reflect the dimension INNOVATIVE/CONSERVATIVE, this would suggest

778 a tendency of YUM in content questions to semantically bleach out and, perhaps while carrying

779 sentence-final question prosody,[34] develop into a general content question particle. Such

780 considerations are irrelevant for polar questions, since the polar question clitic =*UU* is consistently

781 used.

782

783 **7. YUM in connectors**

784

785 Apart from its uses at the end of matrix clauses, reflexes of *jaʊma* form part of four clausal

786 connectors: conditional YUM *bol* 'if', similative YUM *šig* 'as if', causal *yumčin'* 'since' and

787 alternative *yumuu* 'or'. Syntactically, YUM *bol* and *yumčin'* function as subordinating conjunctions,

---

[32] Judging from the not entirely complete annotation of Brosig (2015, see 47-48: fn 3), SC contains 99 structural participle-based content questions without YUM, 68 of which contain a question particle and 10 more are embedded by the quotative verb *ge*-. Of the remaining 21 questions, 12 are part of enumerations, topicalizing questions that the speaker herself immediately answers etc., leaving only 9 instances in which a participle (in all 9 cases the perfect participle *-saŋ*) forms a syntactically free-standing question without question particle.

[33] Since IC is not annotated for word class and semi-automatic morpheme-string-based annotation is not feasible due to the large set of false-positives, we operationalized the sequence "interrogative pro-word + word + YUM" with the interrogative pro-words *xezee* 'when', *yaaɢaad* 'why', *yaaž* 'how', *xed* 'how many', *xaan* 'where', *xeŋ* 'who', *yuu* 'what' (for the last three including case forms). Participles were defined as *-dAg / -sAŋ / -x / -AA*. For SC, this method would have identified 42% of relevant sequences. If so, this would mean that content questions containing YUM are 28.6 times more frequent in SC than in IC. Since SC is very interview-heavy, while IC contains many newspaper articles and similar narrative text types, this is conceivable.

[34] Karlsson (2003) observes a rising fundamental frequency as a general option consistently realized by 2 of 7 speakers for all question types in her sample of isolated read sentences.

788   while the versatile, somewhat postposition-like YUM *šig* serves, for example, as complementizer

789   for morphologically passive cognitive/perceptive verbs, and *yumuu* coordinates clauses or even

790   phrases. All four forms also allow for sentence-final uses. However, both *yumčin'* and *yumuu* are

791   lexicalizations in which the morph *yum* cannot be cliticized to the preceding word, and both

792   exhibit more or less discourse-connective semantics. *Yumčin'* is presuppositional, but trivially so

793   due to its causative semantics. *Yumuu*, in turn, which must have developed from alternative

794   questions (*A yum=uu, B yum=uu?* 'Is it A or is it B?'), has synchronically lost its original presumptive

795   or epistemic meaning, even though its etymology might still render it more inclusive (in the sense

796   'or also' as opposed to 'either … or') than alternative forms such as *eswel* 'or' (< *ese a-bala 'if it

797   is not'). With the conditional (YUM) *bol* and similative (YUM) *šig*, in contrast, the presence or

798   absence of the clitic YUM is usually semantically rather than structurally conditioned. An

799   exception to this is the use of YUM *šig* with adjectives, in which case YUM functions as a

800   semantically empty nominalizer for some informants. In other contexts, however, YUM functions

801   as the assertive particle, which may implicate notions ranging from (the speaker's) presumption

802   to (agreed-upon) certainty (i.e. presupposition). YUM thus indicates a gradual increase in the level

803   of subjective epistemic commitment as compared to bare conditionals or similatives. In the

804   following, we will discuss YUM *bol* (§7.1) and YUM *šig* (§7.2), while leaving aside the lexicalized

805   forms *yumčin'* and *yumuu*.

806

807   *7.1 YUM in conditionals*

808

809   Conditional clauses are regularly formed with the conditional converbal suffix *-WAl* or, if relative

810   tense and some other features have to be marked, by the etymologically related conditional

811   particle *bol*. YUM *bol*, in turn, denotes presumptive conditional clauses, in which the speaker is

812   strongly inclined to believe that the condition actually holds. This view is shared by Mönx-

813   Amɢalaŋ (1996: 43) who describes YUM *bol* as emphasizing causation over conditionality. The

814   difference between the two forms is illustrated in (34a-c):

815

816   (39)(a) *LYRICS üg* **buruu bič-seŋ** **bol** *uučlalt xüs-ii!*[35]

817   lyrics   word wrong   write-PTCP.PRF if   excuse   wish-VOL

818   {Statement made while posting a video with text on an internet video platform:}

819   'If I should have written down the lyrics incorrectly, I beg your excuse.'

820   (b) A: *end aldaa ɢar-čee*.

821   here mistake come.up-PST.IMM.INFER

822   [admits, in reaction to B's complaint] 'A mistake has occurred here.'

823   B: **buruu bič-seŋ** **yum bol** *ter-iig=ee* *zas-aač*[36]

824   wrong   write-PTCP.PRF   ASS   if   that-ACC=RPOSS   correct-IMP(imploring)

825   'If you [indeed agree that you] have written [it] incorrectly, please correct it!'

826   (c) *Eŋgiiŋ üg-eer bič-seŋ* **bol** *deer bai-ž.*[37]

827   normal   word-INS write-PTCP.PRF   if   above   AUX-PST.IMM.INFER

828   'If I had written it in plain words, it would have been better.'

829

830   In (34a), where YUM is absent, it is merely considered possible that mistakes occurred. In (34b)

831   with YUM, the interlocutors have already reached agreement on this point, and the condition is

832   indeed presupposed. Without this previous agreement, the occurrence of mistakes would merely

---

[35] https://www.youtube.com/watch?v=GW8oxnbhleI, 2009-03-19, retrieved 2017-02-23
[36] http://forum.asuult.net/viewtopic.php?t=78968, 2005-12-06, retrieved 2017-02-23
[37] https://www.mongolianeconomy.mn/en/p/6912, 2014-11-04, retrieved 2017-02-23

833 be presumed, i.e. considered as likely by the speaker alone. In irrealis conditionals, like (34c), non-

834 referential YUM must be absent.

835   Apart from its subordinating clause-final use, *bol* also occurs in sentence-final position in two

836 functions. Firstly, it can introduce conditions as afterthoughts, which leads to the same semantic

837 patterns just discussed. Secondly, it may be used as a question particle for content questions

838 which then expresses that the speaker is puzzled and wondering about the issue at hand, and

839 might then ask for approximate rather than precise answers. In this context, the distinction

840 between the presence and absence of YUM parallels its presence and absence in regular questions

841 as discussed in §6.2, i.e. between presumptive questions as in (40) and actual questions as in (41):

842

843 (40)*Soc'aalizm-iiŋ üy-d=č   ma(n)-ai-x biy_daasan tusɢaar_togtnosoŋ, büreŋ erx-t*

844   socialism-GEN   period-DAT=FOC 1PL-GEN-NMLZ autonomous   independent   full   right-COM

845   *uls bai-sŋ-iig   oloŋ žil "ax   düü-(g)iiŋ   nairamdalt xar'alcaa-tai"*

846   state AUX-PTCP.PRF-ACC many year older.brother younger.brother-GEN amiable   relationship-COM

847   *yaw-saŋ "ax nar"=maan' yaaɢaad **med-ee=güi   öŋgör-sön yum bol=oo**?* (IC)

848   go-PTCP.PRF older.brother-PL=STC why   realize+know-PTCP.RES=NEG pass-PTCP.PRF ASS   if=DP

849   'Even at the time of socialism, why <u>did</u> our "elder brothers" [from Russia] who lived in an

850   "amiable relationship of older and younger brother" with us, <u>ignore</u> for so many years

851   that ours was an autonomous, independent, plenipotentiary state?'

852 (41)*En üzegdel   xer udaaŋ **ürgelžl-ex   bol=oo**.* (IC)

853   this phenomenon how long   continue-PTCP.NPST if=DP

854   'How long <u>might</u> this phenomenon [i.e. having such kind of sexual urges] <u>last</u>?'

855

856   In (40) with YUM, a national-conservative politician puts forward a rhetorical question in order to

857   suggest a "chauvinist, contemptuous" stance of Soviet Russia towards Mongolia. In YUM-less (41),

858   the writer who just detailed her socially slightly non-conforming sexual preferences is sincerely

859   asking for advice.

860

861   *7.2 YUM* in similative complements

862

863   Similiative constructions are based on the distributionally unique, somewhat postposition-like

864   word *šig* 'like', which under certain conditions may or has to be preceded by YUM. Depending on

865   construction type and the word class preceding it, the function of YUM differs. It either

866   nominalizes adjectives or disambiguates strong epistemic probability from actual resemblance. In

867   the following, we will look at the sequence "NOUN/ADJECTIVE/PARTICIPLE (YUM) *šig*" as part of three

868   construction types: similative attributes, simulative identity predications, and complements of

869   passive perception verbs.

870   If *šig* introduces similative attributes, e.g. to *xüŋ* 'person', the clitic YUM never precedes it.

871   When the collocation *yum šig xüŋ* is attested (n_{IC}=7), *yum* is a person-referring noun as in (42):

872

873   (42)*Bi=čin' dar'-tai   torxoŋ deer suu-ž bai-(ɢ)aa **yum šig xüŋ**   šüü dee*. (IC)

874   1SG=STC   powder-COM keg   on   sit-CVB AUX-PTCP.PRS thing like person DP   DP

875   'I am <u>somebody like the (poor) thing</u> sitting on the powder keg.'

876

877   In similative identity constructions of the structure (YUM) *šig bai-n* NMLZ/ASS like AUX-PRS.IMM and

878   passive perception constructions of the structure (YUM) *šig sana-gd-* NMLZ/ASS like seem(think.of-PASS-),

879   speakers rarely use bare adjectives (e.g. $n_{IC(SIMILATIVE.IDENTITY)}$=4/280 ~ 1.4% and $n_{IC(PASSIVE.PERCEPTION)}$=24/681

880 ~ 3.5%, respectively) or any words (including nouns) to which a focus clitic is attached. Instead,

881 speakers usually combine these adjectives with other elements, typically YUM and occasionally

882 *baix*. In the similative identity construction in (43) and the passive perception construction in (44),

883 YUM is thus understood by some informants as a semantically empty, but structurally required

884 nominalizer. With words that are marked by the limitative focus clitic =*l* or the additive focus clitic

885 =*č* as in (45), there is indeed no alternative to YUM (which is almost always present, e.g. in

886 $n_{IC(PASSIVE.PERCEPTION)}$=207/209).[38] However, for adjectives there is a certain dichotomy between YUM

887 and the non-past participle copular auxiliary form *bai-x* (10.7%), which also can be interpreted in

888 semantic terms.[39] Thus, while YUM in (43) is understood by some informants to convey the

889 speaker's commitment to a guess, *baix* in (46) rather conveys a proposition that is posited on the

890 basis of visual evidence available at the time of speech. Given a suitable conversational context,

891 these informants would accept either YUM or *baix* in both (43) and (46).[40]

892

893 (43)*Dawkaa bid xoyor ix* **tös(-)tei** *yum šig bai-n*. (IC)

894 NAME 1PL two very similar(similarity-COM) thing/ASS like AUX-PRS.IMM

895 'Davka and I seem to be very similar (people).'

896 (44)*bi=č_geseŋ žiremseŋ bai-saŋ. bas arai=l* **ert yum** *šig sanagda-ž*

897 1SG=even pregnant AUX-PST.EST also rather=FOC early NMLZ/ASS like seem-CVB

898 **bai-saŋ=č** *odoo saixaŋ bai-n*. (IC)

899 AUX-PTCP.PRF=FOC(>although) now beautiful AUX-PRS.IMM

---

[38] Both exceptions involve the non-past participle -*x*.

[39] *Baix* means 'is' if used attributively. While in sentence-final position it historically meant 'will be', but its synchronic meaning in the latter position is 'maybe'.

[40] The nature and size of our informant sample (see §2) does not allow any conclusion about potential social variables that could account for this variation.

900   'I, too, was pregnant. Although it seemed rather early to me, too, now it feels wonderful.'

901   (lit. '... like an early thing ...')

902   (45)*ter   neeree šal   oncgüi   zaluu=č   yum šig sanagd-čix-laa*. (IC)

903   DEM.DIST indeed   completely unpleasant young.man=FOC NMLZ like seem-COMPL-PST.IMM

904   'He indeed (inadvertently) seemed like a very unpleasant fellow to me.'

905   (46)*Coŋx-oor xara-xad   gadaa **ix goy bai-x   šig bai-n***. (IC)

906   window-INS look-CVB.when outside very nice AUX-PTCP.NPST like   AUX-PRS.IMM

907   'When I look through the window, it seems to be very nice outside.'

908

909   However, if the preceding element is a noun, the presence of YUM is not uncommon in the passive

910   perception construction (n_{IC}=118/528, 22.3%) and is even frequent in the similative identity

911   construction (n_{IC}=127/242, 52.5%). The basic distinction between zero and YUM in this type of

912   context thus is very evidently semantic rather than structural, which we shall illustrate with

913   examples from the similative identity construction. Zero in (47) expresses the proposition that

914   two entities resemble each other: the parliamentarians in question are adults, but resemble

915   children in sharing some of their traits. In contrast, YUM in (48) expresses epistemic uncertainty

916   on the part of the speaker: the speaker contemplates that the subject is (and does not merely

917   resemble) a cute girl, but is not sure as to whether this is actually true. The absence of YUM would

918   report an observation of actual resemblance on the basis of visual evidence.

919

920   (47)*Gišüü-d-iig   **xüüxed šig bai-n**   ge-x   yum*. (IC)

921   member-PL-ACC child   like AUX-PRS.IMM say/COMP-PTCP.NPST ASS

922   'One might say the members [of parliament] are like children.'

923   not: 'One might say the members [of parliament] seem to be [actual] children.'

924   (48)*Bod-wol*   ***ix***   ***xöörxöŋ=l ox'oŋ*** ***yum šig bai-n.*** (IC)

925   think-CVB.COND very cute=FOC   girl   ASS   like AUX-PRS.IMM

926   '(But why doesn't she find anything but money to think about?) If you look at her, she just

927   <u>seems to be a very cute girl</u>.'

928

929 Apart from nouns, adjectives and words marked by focus clitics, the two constructions under

930 discussion may involve participles as a fourth major morphological class. Here, an actual

931 resemblance between subjects (which typically denote physical entities) and the events coded by

932 these participles is usually impossible, so the clear-cut semantic distinction found with nouns

933 cannot be at work here. The frequency of YUM with non-focused participles of all verbs except

934 *bai-* in the similative identity construction is at 92.1% and might be approaching the threshold of

935 structural obligatoriness (n_{IC}=741/804), but this can be ruled out for the passive perception

936 construction in which YUM only collocates with 78.8% (n_{IC}=781/991) of such participles. Similar

937 to the contrast between (43) and (46), the contrast here runs between confident guessing as

938 indicated by YUM in (49) and objective (if approximate and incomplete) recollection as indicated

939 by the absence of YUM in (50). For some informants, guessing implies the absence of actual

940 firsthand knowledge, so for them the presence of YUM in combination with the contextually given

941 first-person cognitor would render (50) unacceptable. In both examples, one could use *baix*

942 instead of either YUM or zero to indicate that the speaker is interpreting visual evidence that is

943 immediately accessible to her at the time of speaking.

944

945   (49)*90-eed*   *oŋ-oos ömön manai šüüx, cagdaa xariŋ=č ɢaigüi*

946   1990-COLLECTIVE   year-ABL before 1PL.GEN court police   but=FOC   passable

947   *šudrɑɢ* ***ažilla-dag*** ***bai-saŋ***   ***yum šig sana-gd-dag***. (IC)

948   upright   work-PTCP.HAB   AUX-PTCP.PRF   ASS   like   think.of-PASS-PTCP.HAB

949   'But it seems to me that before the 1990s our courts and police were indeed doing their

950   work rather honestly.'

951   (50)*Tegeed ter möŋg-öör=öö nair*   **xii-ž bai-saŋ**   **šig sana-gd-(a)ž bai-n**. (IC)[41]

952   then   that money-INS=RPOSS   celebration do-CVB   AUX-PTCP.PRF like   think.of-PASS-CVB AUX-PRS.IMM

953   '[Along with the title of "distinguished [sportsperson]", they were at that time awarding

954   me 2000 tugrik.] I recall that with that money we were holding a celebration.'

955

956 Judging from this data, it appears that YUM in similative complements with an opposition

957 between zero, YUM and *baix* expresses a subordinated guarded assertion that is supported by

958 evidence other than observation, but subjectively supported by the speaker. The extent to which

959 such variation exists is not entirely clear, though, and requires further research.

960

961 **8. The presumptive reasoning construction YUM + copula**

962

963 Next to sentence-final (§6) and (quasi-)clause-final uses (i.e. those discussed in §7), YUM also

964 combines with the copula mainly to express presumptive reasoning, i.e. a form of inference that

965 cannot be straightforwardly deducted from concrete observations but involves some defeasible

966 general considerations that the speaker commits to. The relevant construction consists of a

967 nominal clause followed by YUM, the copula *bai-*, and an ongoing present tense suffix (-*n* or -*AA*).

968 Möŋx-Amɢalaŋ (1996: 42, our translation) describes the main function of *yum bai-n* as follows:

969

---

[41] Р.Даваадалай: Биднийг гавьяа байгуулаад ирэхэд баярын бичиг ч өгөөгүй,
http://davka0624.blog.gogo.mn/read/entry192001, 2010-12-16, retrieved 2017-04-03

970   The event passed when [the speaker] was not aware of it. But given this, even though the

971   speaker has insufficient information about the event and is connected to it only indirectly,

972   s/he relates to the event by drawing some form of conclusion (...).

973

974 This description captures a process of inference that is not based on immediate visible evidence,

975 but on general knowledge or other kinds of information (cf. Aikhenvald 2004: 2-3) that don't allow

976 for proper (near-infallible) inference. This indirect reasoning process is illustrated by (51) and (52):

977

978   (51)*tan-ai kompan' maš saiŋ bar'alaɢ* **bar'-ž**

979   2GEN--PL company very good construction construct-CVB

980   ***bai-(ɢ)aa*** **yum** ***bai-n***. (IC)

981   AUX-PTCP.PRS ASS/NMLZ AUX-PRS.IMM

982   [I told C, "Take a look at the quality of our construction work", and showed it to him. C

983   then said:] "Your company is apparently doing construction work very well. [I will help

984   you. Let me find out what possibilities there are]." (Brosig & Skribnik 2018: 563)

985   (52)(a) *bi yag=(a)l öwčöŋ* **aw-saŋ** **yum** ***bai-n*** (IC)

986   1SG precise=FOC illness take-PTCP.PRF ASS/NMLZ AUX-PRS.IMM

987   ('My genital emitted a bad smell, and white liquid came out of it amass, and it was

988   itching. I thought) "I've precisely contracted a disease."'

989   (b) *belg-iiŋ zam-iiŋ xaldwart öwčöŋ* **aw-saŋ** ***bai-n***.[42]

990   genital-GEN way-GEN infectious illness take-PTCP.PRF AUX-PRS.IMM

---

[42] www.cekc.mn/2015/04/Хэлэхгүйгээр-нууцаар-эмчлэх-байсан-г, 2015-4-15, retrieved 2016-12-13

991   '(I couldn't tell it to anyone and just went directly to the women's hospital and showed

992   it to them. Upon seeing it, they said:) "You've contracted a sexually transmitted

993   disease."'

994

995 In (51), the speaker has some firsthand knowledge of the construction site, but is still drawing an

996 evaluative conclusion about a present situation from multiple connected observations after the

997 fact. The contrast between presence and absence of YUM is illustrated in (52). In (52a) [with YUM],

998 a layperson infers that he has acquired an illness from symptoms of which he only has a vague

999 understanding. In (52b) [a present perfect without YUM], a medical doctor evaluates the same

1000 kind of evidence on the basis of her/his scientific knowledge. While the speakers of (51) and (52a)

1001 highlight that they arrived at the proposition through a subjective reasoning process that drew

1002 from incomplete information, they don't necessarily relativize it in terms of factuality or

1003 subjective certainty. This can be seen from the meaning of typical matrix verbs with which the

1004 presumptive reasoning construction occurs (see Table 4):

1005

1006 *[intended location of Table 4]*

1007

1008 Here, the two most common matrix verbs *bod-* and *oilɢ-*, as in (53), refer to a reality that has been

1009 arrived at through the speaker's reasoning processes, but that nevertheless is assessed as

1010 objectively true. Among the three next-common forms, two indicate the speaker's uncertainty

1011 about reality, either in terms of non-commitment (*bod-ogd-* think-PASS-) or partial commitment

1012 (*dügn-*), and one (*üz-*) commits the speaker, but doesn't lay claim to objective reality. While the

1013 meaning of YUM *bai-n* thus still appears to be somewhat compatible with matrix verbs that

1014 relativize speaker commitment or the speaker's claim to objectivity like *bod-ogd-*, *dügn-* and *üz-*,

1015 it mostly co-occurs with matrix verbs that commit the speaker like *bod-* and *oilɢ-*, which in turn

1016 suggests that the use of YUM *bai-n* itself tends to express such commitment.[43]

1017

1018 (53)*Ter üy-d   en büx(n)-iig xii-x   xüsel zor'oɢ, ermelzel čadwar MAXN,*

1019   that time-DAT this all-ACC   do-PTCP.NPST wish   courage aspiration ability   PARTY

1020   *AN-d   bai-x=güi   yum   bai-n   ge-ž   oiloɢ-soŋ.* (IC)

1021   PARTY-DAT AUX-PTCP.NPST=NEG ASS/NMLZ AUX-PRS.IMM COMP-CVB understand-PST.EST

1022   'At that time we <u>understood</u> that the wish and courage, aspiration and ability to

1023   undertake all of this <u>is absent</u> in the Mongolian People's Revolutionary Party and the

1024   Democratic Party. (Therefore, we considered it correct to establish a separate party.)'

1025

1026 The presumptive reasoning construction encodes event time and evaluation time separately. As

1027 can be seen by comparing ongoing present forms such as *bar'-ž bai-ɢaa* 'is constructing' in (51)

1028 and *bai-x=güi* in (53) with the perfect *aw-saŋ* 'has taken' in (48a), the time of the event itself is

1029 expressed by the (formally attributive) participial forms that precede YUM. The time at which the

1030 evaluation is drawn and holds true is expressed through the tense form of the copula. It can only

1031 be an ongoing present form, as the use of other forms such as the Established Past in (54) entails

1032 the interpretation of YUM as a noun. If the Immediate Present *-n* is used (n_{SC}=115) as in (51)-(53),

1033 it involves the speaker assessing the event at the time of speech. If the Established Present *-AA*

1034 used (n_{SC}=20), it refers to an inference that the speaker drew earlier and still advocates, but this

---

[43] One could also argue (as one of the reviewers did) that overall frequency determines the likeliness of a given lexical item to be used in a particular construction. From this perspective, any given token of *bodogd-* relative to its overall frequency in IC is more than twice as likely to be used in this particular construction than any given token of *bod-* or *oilɢ-*, while while *dügn-* would be half and *üz-* (due to the relatively low text frequency of its cognitive extension) less than a tenth as likely. While this observation is certainly relevant, we still maintain that the meaning of a particular construction is determined by absolute rather than relative frequencies.

1035 collocation is usually accompanied by a modal particle such as *biz* and used in contexts where the

1036 speaker wants to cross-check this inference with the addressee, as illustrated in (55).

1037

1038   (54)*ergelz-(e)x*   ***yum*** *bai-saŋ*   *biz=dee*. (IC)

1039   doubt-PTCP.NPST   thing   AUX-PST.EST   DP=DP

1040   'There were <u>things</u> to be skeptical about, right?'

1041   (55)*yuu=g-saŋ*   *üg=üü*   *öör=öö, öör-iiŋ*   *eež*   *aaw*

1042   what=say-PTCP.PRF word=Q   self-RPOSS self-GEN   mother father

1043   *xoyor* ***bai-dg=iin***   ***bai-(ɢ)aa***   *biz=dee*. (SC)

1044   two   be-PTCP.HAB=ASS/NMLZ   AUX-PRS.EST   MP=DP

1045   'What are you saying there? Your parents also <u>live at your place</u>, don't they?'

1046

1047 The presumptive reasoning construction can be used with first person participants who recollect

1048 with effort (cf. Möŋx-Amɢalaŋ 1996: 42) witnessed events as in (56). With regular presumptive

1049 reasoning, "effortful recollection" shares the conscious, effortful thought process that aims at

1050 inferring some state-of-affairs from unreliable evidence (general knowledge, vague memories)

1051 and the somewhat reduced reliability of the conclusion ultimately arrived at.

1052

1053   (56)***Bi*** *döröw, tawaŋ žil-iiŋ ur'd neg* ***unš-saŋ***   ***yum***   ***bai-n***. (Möŋx-Amɢalaŋ 1996: 42)

1054   1SG four   five   year-GEN before one   read-PTCP.PRF ASS/NMLZ AUX-PRS.IMM

1055   'I <u>remember that I read</u> this once four, five years ago.'

1056

1057 As can be seen from (54), the presumptive reasoning construction is structurally similar to the

1058 existential construction. While ambiguity cannot usually be assumed for most participle-based

1059   presumptive sentences, referential interpretations such as 'I am precisely <u>a poor thing</u> who

1060   contracted a disease' for (48a) or 'The wish (...) to (...) was <u>something</u> absent in the Mongolian

1061   People's Revolutionary Party (...)' for (53) are theoretically conceivable with unreduced accented

1062   *yum*. Actual cases of ambiguity (in reading) are mostly found with adjectival predicates as in (57),

1063   for which a reader, without any prosodic information or contextual disambiguation, might indeed

1064   assume a presumptive interpretation as in (53a) or a referential interpretation as in (53b).

1065   Ambiguity does not usually arise with nouns for which attributive interpretations are much less

1066   accessible; the interpretation for YUM *bai-n* (along with YUM *bol*) in (58) is decidedly non-

1067   referential. Ambiguity does not arise with numerals, quantifiers or pro-adjectives (like *iim*

1068   'such$_{proximal}$') either, but for the opposite reason: since they strongly favour attributive

1069   interpretations, they always yield a referential interpretation as seen in (59).

1070

1071   (57)*gexdee*   *aygüi* ***zöölöŋ yum***   ***bai-n****. (SC)

1072   but   fairly   soft   thing/NMLZ   AUX-PRS.IMM

1073   (a) 'But it apparently turns out to be fairly soft.'

1074   (b) 'But it is a fairly soft thing.'

1075   (58)*xulɢaič yum bol ter* *ünexeer* ***soŋɢodog xulɢaič yum bai-n****. (IC)

1076   thief   ASS   if   that   really   exceptional   thief   ASS   AUX-PRS.EST

1077   '(But this time there is nobody who matches [former president] Enkhbayar.) If he is indeed

1078   a thief, then he is really an exquisite thief / a really exquisite thief.'

1079   (59)*bi*   *bai-wal*   ***büx***   ***yum*** *bai-n*. (IC)

1080   1SG   be-CVB.COND   all   thing   be-PRS.EST

1081   'If I am there, <u>everything </u>is there.'

1082

1083   Structurally, the presumptive reasoning construction resembles the action denial construction in

1084   (60) which usually features an overt or covert first-person possessor and takes *bai-x=güi* AUX-

1085   PTCP.NPST=NEG or *alaɢ* 'absent' as main predicate. The possessor can be nominative, in which case it

1086   parallels a privative possessive pattern (cf. Brosig 2015b: 98), where the referential interpretation

1087   of YUM (60b) is synchronically no longer acceptable. Alternatively, it could be in the dative, in

1088   which case it would parallel a negated locative possessive construction (see Brosig 2015b: 100-

1089   101). Similarly to the presumptive reasoning construction, the actual synchronic meaning (60a)

1090   suggests a flat clause structure which could in theory also go back to a subject that is internal to

1091   the relative clause as suggested by (60c), but such an attributive clause (here and in the

1092   presumptive reasoning construction) should allow for an internal genitive subject (cf. Seesing

1093   2013 on Kalmyk). However, genitive subjects are already impossible in the pattern PARTICIPLE (*yum)*

1094   *biš* 'it is not so that …, but rather …' (Mukai 2001: 76-78), and the same holds for the presumptive

1095   reasoning and action denial constructions.[44]

1096

1097   (60)*bügd=l   biy=ee   ünel-d=ii   šüü yag   yas yum-an deer=ee …*

1098   all=FOC   body=RPOSS put.a.price.on-PTCP.HAB=ASS DP   precisely bone thing-ATT on=RPOSS

1099   *bi law   neeɢ üzeŋ_yad-aad bai-x   yum alaɢ … (IC)*

1100   1SG certainly very   hate-CVB.PFV   AUX-PTCP.NPST NMLZabsent

1101   (a) 'All [people] are prostituting themselves, essentially at the heart of the matter … I

1102   definitely don't hate this particularly much …'

1103   (b) [impossible:] 'I certainly have <u>nothing</u> that I keep on hating particularly much.'

---

[44] Mukai (2001: 76-78) observes that genitive subjects are likewise impossible in the pattern PARTICIPLE (*yum*) *biš* 'it is not so that …, but rather …'. However, this only shows that *yum* doesn't behave as a noun in this context, which, given that it is facultative in this context, is synchronically to be expected. Semantically, Jīngāng (2007: 15) interprets YUM in YUM *biš* as denying that an assertion is based on a more general understanding, i.e. in the same way as sentence-final YUM in declaratives. In this respect, it resembles YUM in structurally non-obligatory contexts of §7.

1104   (c) 'There is no respect in which I keep on hating [this] particularly much.'

1105

1106   As far as synchronic structure is concerned, we may at any rate conclude that the collocation YUM

1107   *bai-n/-(ɢ)aa* does not structurally subordinate, but that it fulfills a copula-like function. At the

1108   same time, semantically, YUM *bai-n* brings about a presumptive meaning similar to YUM in

1109   questions and at the end of subordinate clauses. Arguably, it might do so by combining the

1110   speaker's subjective support for a proposition as expressed by sentence-final YUM with the

1111   perceptive evidence that is suggested by *-n* in *bai-n.*

1112

1113   **9. Discussion and conclusions**

1114

1115   In this paper, we have analyzed the synchronic uses of the particle/clitic YUM, an erstwhile

1116   nominalizer that has developed from the noun *yum* '(some)thing' in Khalkha Mongolian. In the

1117   sentence-final position of declaratives, YUM expresses the speaker's strong subjective

1118   commitment to a proposition, thus increasing its assertive force. In interrogatives, the speaker

1119   commits by presuming a particular answer, which in sincere questions and guesses can be

1120   interpreted as anticipation and in exam questions as a claim to the speaker's full knowledgeability.

1121   Since agreement between the interlocutors on a certain state-of-affairs is anticipated in rhetorical

1122   questions and exclamatives, the proposition is even presupposed here. Underdeterminacy

1123   between subjective commitment, anticipation and a claim to actual knowledge is also found in

1124   noun- and participle-based conditionals and similatives. If combined with a marker of direct

1125   perception at speech time (of an event itself or of evidence for that event), the presumptive

1126   reasoning construction combines the inferential dimension of the latter with the subjective

1127   dimension of the former, thus bringing about the meaning of the speaker's personal

1128   interpretation of the evidence on the basis of objective external evidence. Overall, subjective

1129   commitment or assertive force in finite declaratives and presumption, presupposition or

1130   epistemic strength in non-declarative contexts seem to be conceptualized as different sides of the

1131   same coin.[45]

1132   In spite of the assertive and epistemic dimensions that YUM has gained, it has not fully

1133   ceased to function as what can most likely be assumed to be its source function, namely as a

1134   general noun that can be used as a dummy for other word classes in constructions that require

1135   the use of nouns. For instance, YUM does not equally contribute assertive/presumptive meaning

1136   when used in adjectival complements of similative constructions or in exclamatives, since its

1137   presence in these constructions is, to a specifiable extent, structurally required.[46]

1138   Into the system of Khalkha Mongolian assertions, the forms YUM and YUM *bai-n* introduce

1139   an important distinction not found in most other Central Mongolic dialects. For instance, among

1140   the past tense forms of Khalkha Mongolian, all forms except for the quotative *-san ge-*

1141   unequivocally and without epistemic modal graduation assert a certain state-of-affairs, but the

1142   epistemological foundation for asserting differs, as seen in Table 5.

1143

1144   *[intended location of Table 5]*

1145

1146   In this analysis, all forms except *-saŋ ge-* assert a certain state-of-affairs, the only two forms that

1147   personally commit the speaker, to a varying degree, to the truth of the assertion are *-saŋ* YUM

1148   and *-saŋ* YUM *bai-n*. But even here, the commitment is not meant as a claim of the speaker's

---

[45] Presupposing, due to its intersubjective epistemological component, defeasibly implicates a stronger epistemological commitment of the speaker than presuming, i.e. implicating certainty rather than strong probability.

[46] Möŋx-Amɢalaŋ (1996: 44) argues along similar lines for YUM *bol*, but his examples are not chosen well.

1149 infallibility or of objective truth, but merely as an indication of the status that said information

1150 occupies within the speaker's representation of the world, thus enabling the addressee to

1151 independently assess its reliability. In contrast, if a speaker were to claim something resembling

1152 objective truth (which is in itself a culture-specific concept [cf. Hughes 1988: 61-62 cited in

1153 Wierzbicka 2003: 103]), she would have to resort to discourse markers such as the compound

1154 particle *šdee*[47] which presents the speaker's knowledge as superior to the addressee's (*šüü*) and

1155 indicates that reality has to be constructed in this particular way on grounds that are independent

1156 of what the interlocutors might personally wish (*dee*). Pragmatically, though, YUM does often

1157 implicate the speaker's wish to convince the addressee of the speaker's claim, accommodating

1158 for the Gricean perspective, while bare participial and especially finite forms are much more

1159 suitable for presenting information to the addressee on a take-it-or-leave-it basis (as proposed by

1160 Dummett 1981). For a competent adult speaker to assert, at any rate, proves to be something so

1161 firmly entrenched in language-specific epistemological systems that any valid cross-linguistic

1162 generalizations can only be derived from a typological basis. The social consequences that a false

1163 statement (uttered using a particular linguistic form including those containing YUM) might entail

1164 for the speaker, in turn, are an empirical question that could perhaps best be studied on the basis

1165 of speaker intuition as instantiated in truth-focused discourse such as arguments or court hearings

1166 and its lexical codification in word fields including terms such as *xudlaa* 'lie, false, fake' which

1167 seems to cover a much wider, but as of yet not precisely delimited semantic range than its closest

1168 English translation equivalents.

1169 This paper must leave multiple tasks for future research. One relevant issue is the

1170 grammaticalization of *yum*. This includes both the uses discussed in this paper (for which it would

---

[47] Or to the modal particle *biz* the *z* of which is cognate with a Middle Mongol discourse particle *je* which requires more research but seems to have expressed either high probability/necessity or the speaker's conviction (cf. Brosig 2014: 14).

1171   be particularly helpful to find out whether finite or non-finite uses developed first) and the

1172   secondary development of *yum* as a functional noun into the discourse markers *yumuu* 'or' and

1173   *yumčin* 'since'. Secondly, for want of space, we did not discuss the meanings that *yum* expresses

1174   in connection with the irrealis particle *san* that relate to unfulfilled wishes or futile aspirations.[48]

1175   Thirdly and finally, while the current analysis roughly identifies the meaning and functional range

1176   of YUM, it does not predict when it is infelicitous to use YUM, an issue that is most problematic

1177   for finite declaratives which might benefit from a discourse-analytic approach.

1178

1179   **Glosses**

1180

1181   Below we list all glosses that are not part of the Leipzig Glossing Rules.

1182

1183   *[intended location of Table with glosses]*

1184

1185   **Bibliography**

1186

1187   Aikhenvald, Alexandra & Robert Dixon (eds.). 2014. *The grammar of knowledge: A cross-linguistic*

1188   *typology*. Oxford: Oxford University Press.

1189   Bach, Kent & Robert Harnish. 1979. *Linguistic communication and speech acts*. Cambridge, MA:

1190   MIT Press.

---

[48] For instance, *Ter ner_dewšigč-d-ees bi Amaržargal bagš-iig yal-aasai ge-ž ix zalb'ar-saŋ yum-saŋ* (IC) that candidate-PL-ABL 1SG NAME teacher-ACC win-DES QI-CVB much pray-EST.PST ASS-IRR 'Among the candidates, I prayed that Amarjargal, our revered teacher, would win. (But it didn't come about.)'

1191 Barðdal, Jóhanna. 2011. The rise of dative substitution in the history of Icelandic: A diachronic

1192   construction grammar account. *Lingua* 121(1). 60-79.

1193 Bayancoɣtu. 2002. *Qorcin aman ayalɣun-u sudulul* [A study of the Khorchin dialect]. Kökeqota:

1194   Öbür mongɣul-un yeke surɣaɣuli-yin keblel-ün qoriy-a.

1195 Behrens, Leila. 2012. Evidentiality, modality, focus and other puzzles. In Andrea Schalley (ed.),

1196   *Practical theories and empirical practice: A linguistic perspective*, 185-244. Amsterdam:

1197   Benjamins.

1198 Binnick, Robert. 2012. *The past tenses of the Mongolian verb*. Leiden: Brill.

1199 Bläsing, Uwe. 1984. *Die finit indikativischen Verbalformen im Kalmückischen*. Stuttgart: Franz

1200   Steiner.

1201 Brosig, Benjamin. 2012. *Bilee* sul ügiin utaɢ, xereglee [The meaning and function of the particle

1202   *bilee*]. *Xel Zox'ool Sudlal* 5(37). 10-18.

1203 Brosig, Benjamin. 2014. The aspect-evidentiality system of Middle Mongol. *Ural-Altaic Studies*

1204   13(2). 7-38.

1205 Brosig, Benjamin. 2015a. Aspect and epistemic notions in the present tense system of Khalkha

1206   Mongolian. *Acta Linguistica Petropolitana* 11(3). 46-127.

1207 Brosig, Benjamin. 2015b. Negation in Mongolic. *Journal de la Société Finno-Ougrienne* 95. 67-136.

1208 Brosig, Benjamin. 2018. Factual vs. evidential? - The past tense forms of spoken Khalkha

1209   Mongolian. In Ad Foolen, Helen de Hoop & Gijs Mulder (eds.), *Empirical approaches to*

1210   *evidentiality*, 45-75. Amsterdam: Benjamins.

1211 Brosig, Benjamin & Elena Skribnik. 2018. Evidentiality in Mongolic. In Alexandra Aikhenvald (ed.),

1212   *Oxford handbook of Evidentiality*, 554-579. Oxford: Oxford University Press.

1213   Byambasaŋ, P., C. Önörbayaŋ, B. Pürew-Očor, Ž. Sanžaa & C. Žančiwdorž. 1987. *Orčoŋ cagiiŋ*

1214   *moŋɢol xelnii ügzüin baiɢuulalt* [The structure of (verbal) morphology of contemporary

1215   Mongolian]. Ulaaŋbaatar: Šinžlex uxaanii akademi.

1216   Croft, William. 2001. *Radical Construction Grammar*. Cambridge: Cambridge University Press.

1217   de Smet, Hendrik & Jean-Christophe Verstraete. 2006. Coming to terms with subjectivity.

1218   *Cognitive Linguistics* 17(3). 365-392.

1219   Dummett, Michael. 1981. *Frege: Philosophy of language*. 2nd edn. Cambridge, MA: Harvard

1220   University Press.

1221   Goldberg, Adele. 1995. *Constructions*. Chicago: University of Chicago Press.

1222   [Gotō] Goto (後藤), Ksenia. 2009. Sistema finitnych form prošedšego vremeni v kalmyckom jazyke

1223   [The system of finite past tense forms in Kalmyk]. *Acta Linguistica Petropolitana* 5(2). 124-159.

1224   Guntsetseg, Dolgor. 2016. *Differential case marking in Mongolian*. Wiesbaden: Harrassowitz.

1225   Hughes, Geoffrey. 1988. *Words in time: a social history of English vocabulary*. Oxford: Blackwell.

1226   Janhunen, Juha. 2012. *Mongolian*. Amsterdam: Benjamins.

1227   Jary, Mark. 2010. *Assertion*. New York: Palgrave Macmillan.

1228   [Jīngāng] Jingan (金罡). 2007. Mongorugo no shūjoshi yum to nihongo no "da", "no da" nado

1229   bunmatsu keishiki no taishō kenkyū [A contrastive study of the Mongolian sentence-final

1230   particle *yum* and sentence-final forms such as *da* and *no da* in Japanese]. *Shigen (Tōkyō*

1231   *Gaikokugo Daigaku Kijū Gengogaku Ronshū)* 3. 3-21.

1232   [Jīngāng] Jingan. 2010. *Mongorugo no modaritei* [Modality in Mongolian]. Tōkyō: Tōkyō

1233   Gaikokugo Daigaku. Unpublished doctoral thesis.

1234   Johanson, Lars. 2006. Indirective sentence types. *Turkic Languages* 10. 73-89.

1235   Kang, Sin. 2003. Hyen.tay.mong.kol.e.uy yang.thay.chem.sa.ey kwan.han yen.kwu [A study of

1236   modal particles in Modern Mongolian]. *Mongkolhak* 14. 21-52.

1237   Karlsson, Anastasia Mukhanova. 2003. Tonal gestures in Mongolian interrogatives. *Phonum*

1238   *(Reports from the Department of Phonetics, Umeå University)* 9. 189-192.

1239   Kullmann, Rita & Dandii-Yadamiin Tserenpil. 1996. *Mongolian grammar*. Ulaanbaatar: Admon.

1240   Lazard, Gilbert. 1999. Mirativity, evidentiality, mediativity, or other? *Linguistic Typology* 3(1). 91-

1241   110.

1242   [Gāo, Liánhuā] Guu, Lingqu-a (高莲花). 2013. Mongɣul kitad kelen-ü keterkü asaɣuqu ögülebüri-

1243   yi dügümken jergecegülkü ni [A short comparison of rhetorical questions in Mongolian and

1244   Chinese]. *Kele ba Orčiɣulɣ-a* 2013/3. 16-20.

1245   Macagno, Fabrizio. 2018. A dialectical approach to presupposition. *Intercultural Pragmatics* 15(2).

1246   291–313.

1247   Matthews, George. 1965. *Hidatsa syntax.* Den Hague: Mouton.

1248   Michaelis, Laura. 2001. Exclamative constructions. In Martin Haspelmath, Ekkehard König, Wulf

1249   Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals*. Volume

1250   2. Berlin: De Gruyter. 1038-1050.

1251   Mönx-Amɡalaŋ, Yumžiriiŋ. 1996. Moŋgol xelnii "YUM" gedeg ügiiŋ üüreg-utaɢzüiŋ onclog [The

1252   functional and semantic peculiarities of the Mongolian word "ɣum"]. In E. Rawdaŋ & M.

1253   Bazarragčaa (eds.), *Moŋgol yar'aanii xel*, 28-52. Ulaaŋbaatar: Moŋgol Ulsiiŋ Ix Surguul'.

1254   Mukai, Shin-Ichi (向井晋一). 2001. Mongorugo no shōten chōsei keishiki = On the focus of

1255   sentence in Mongolian. *Nihon Mongorugo Gakkai Kiyō = Bulletin of the Japan Association for

1256   Mongolian Studies* 31. 69-90.

1257   Nugteren, Hans. 2011. *Mongolic Phonology and the Qinghai-Gansu Languages*. Utrecht: LOT.

1258   Pagin, Peter. 2016. Assertion. In Edward Zalta (ed.), *Stanford Encyclopedia of Philosophy* (Winter

1259   2016 Edition). https://plato.stanford.edu/archives/win2016/entries/assertion.

1260   Palmer, Frank. 2001[1986]. *Mood and modality*. Cambridge: Cambridge University Press.

1261   Schmidt-Radefeldt, Jürgen. 1977. On so-called 'rhetorical' questions. *Journal of Pragmatics* 1. 375-

1262   392.

1263   Seesing, Olga. 2013. *Die temporalen Infinitkonstruktionen im Kalmückischen*. Wiesbaden:

1264   Harrassowitz.

1265   Song, Jae-mog. 1997. *Tense, aspect and modality in Khalkha Mongolian*. London: University of

1266   London, SOAS. Unpublished doctoral thesis.

1267   Song, Jae-mog. 2002. Grammaticalization of the verb *ge-* 'to say' in Khalkha Mongolian. *Ōsaka*

1268   *Keizai Hōka Daigaku Sōgō Kagaku Kenkyūsho Nenpō* 3. 29-38.

1269   Street, John C. 1963. *Khalkha structure*. Bloomington: Indiana University Press.

1270   Svantesson, Jan-Olof. 1991. *Tense, mood and aspect in Mongolian*. Lund University, Department

1271   of Linguistics, Working Papers 38. 189–204.

1272   Tantucci, Vittorio. 2017. From immediate to extended intersubjectification. *Language and*

1273   *Cognition* 9. 88-120.

1274   Tournadre, Nicolas & LaPolla, Randy. 2014. Towards a new approach to evidentiality. *Linguistics*

1275   *of the Tibeto-Burman Area* 37(2). 240-263.

1276   Umetani, Hiroyuki (梅谷博之). 2013. Classifications of some sentence-final modal particles in

1277   Khalkha Mongolian. *Tokyo University Linguistic Papers* 33. 301-318.

1278   Wierzbicka, Anna. 2003[1991]. *Cross-cultural pragmatics*. 2nd edn. Berlin: de Gruyter.

1279

1280   **Corpora and primary sources**

1281

1282   Baranova, Vlada (created 2007-2014). Narrative corpus of Kalmyk [2 hours, transcribed, glossed

1283   and translated]. Partially published in Vlada Baranova & Sergei Say. 2009. Kommmentarij k

1284   korpusu tekstov & Teksty [Note on texts & Texts]. *Acta Linguistica Petropolitana* 5(2). 710-

1285   852.

1286   Hán, Guójūn (韩国君), Benjamin Brosig, Hasiqimeg, Hasihuu, Mandala, Udbal, Lu Man, Celger.

1287   2012/2016 (created). Corpus of spoken Khorchin Mongolian. 6,5 hours, audio and text,

1288   partially textgrid.

1289   Östling, Robert & Benjamin Brosig. 2011 (created). Corpus of Khalkha Mongolian internet texts.

1290   34,642,000 words, text file.

1291   Saito, Yoshio (斉藤純男). 2008. *The Mongolian Words in the Muqaddimat al-Adab: Romanized*

1292   *text and word index (as of January 2008)*. Tokyo. http://www.u-

1293   gakugei.ac.jp/%7Esaito/Text%20and%20Index.pdf.

1294   Zolžargal, Baasaŋžaw & Benjamin Brosig. 2012 (created). Corpus of unscripted Khalkha Mongolian

1295   TV data. 9 hours, video & text files.