Is There String-Vacuous Verb Raising in Khalkha Mongolian? An Experimental Investigation

Jaehong Shim*

May 5, 2025

1 Introduction

This paper aims to investigate verb raising (henceforth, V-raising) in the Khalkha dialect of the Mongolian language (Mongolic) via scope interpretation between negation and universally quantified arguments.

In syntactic theory, the position of the verb within the syntactic structure has been one of the central issues in the domain of Generative Grammar. Despite cross-linguistic similarities and universality in the verbal category, substantial differences also emerge across languages. One such difference concerns the position of the verb. Beyond its placement relative to core sentence elements such as the subject and object, considerable micro-variation exists even among languages with the same basic word order. For instance, although both English and French exhibit a preferred SVO word order, they differ in a crucial respect: only French shows movement of the verb to a higher syntactic position. Previous studies have proposed a variety of explanations for the driving force behind V-raising (Pollock 1989; Bobaljik and Thráinsson 1998; Koeneman and Zeijlstra 2014). Regardless of theoretical approach, the starting point for these studies has been the surface distribution of verbal elements, as V-raising typically reveals itself in the relative ordering of the verb and other sentence constituents. In other words, you know verbs are raised because you see it.

In this respect, languages with head-final word order—where the verb consistently appears at the right edge of the clause—pose a fundamental challenge for identifying V-raising: the linear position of the verb is fixed. Determining the height of the verb in the syntactic structure becomes difficult, let alone investigating the triggers or motivations for V-raising. As a result, even the existence of the phenomenon itself becomes hard to establish. Consequently, researchers have turned to indirect diagnostics for evidence of V-raising in head-final languages, such as null object constructions (Otani and Whitman 1991; Hoji 1998; Kim 1999), scrambling and coordination (Koisumi 2000; Fukui and Sakai 2003), negative polarity item licensing (Cho and Hong 1988; Choi 1999; Chung and Park 1997; Yoon 1994), coordination of a tensed and untensed conjunct (Yoon 1994),

*I would like to thank Ming Xiang and Erik Zyman for their helpful comments and constructive guidelines for writing up the current paper.

and *do*-support for dislocated affixes (Kuroda 1965; Miyagawa 2001; Choi and Harley 2019). However, due to inconsistencies and conflicting results, the literature has reached no clear consensus on the matter (Han et al. 2007).

The situation is no different in the case of Mongolian. As a head-final language, Mongolian exhibits an even stricter degree of head-finality than other SOV languages discussed in the literature, such as Japanese, Korean, or Turkish, in that it does not permit right-dislocation constructions at all (Guntsetseg 2012, 2016; but see Lee 2023 for evidence of right-dislocation in the Alashan dialect).

This state of affairs raises several questions: where is the verb positioned in the syntactic structure of Mongolian? Is there evidence for V-raising, or does the verb remain in situ? If V-raising exists, is it consistent across speakers, or is there a split in the speaker population, as has been proposed for native Korean speakers (Han et al. 2007)?

The present study aims to address these questions through an experimental investigation. Using native speakers' scope judgments as a diagnostic, I argue that verbs in Mongolian undergo V-raising to the C position.

This paper is structured as follows: Section 2 presents the empirical and theoretical background for the study. Section 3 details the experimental design and results. Section 4 discusses the interpretation of the results and their theoretical implications. Finally, Section 5 concludes the paper.

2 Linguistic background: head-finality and V-raising

This section reviews the issue of V-raising in head-final languages. I will address this issue from several perspectives: the difficulty of detecting V-raising in such languages, the acquisition of V-raising properties, the syntactic properties of Khalkha Mongolian relevant to investigating V-raising, and the conflicting claims in the literature and fieldwork data.

2.1 String Vacuous Verb Raising in Head-final Languages

It is well known that V-raising in head-initial languages can be detected by examining the position of the verb relative to adverbs. For example, Pollock (1989) analyzes the difference in verb–adverb order between English and French as a result of V-raising. In English, post-verbal adverbs are ungrammatical, as shown in (1a), whereas French displays the opposite pattern: the adverb follows the verb, as shown in (1b).

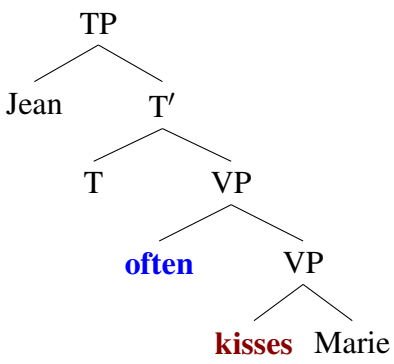
Assuming that adverbs occupy fixed positions within the clausal structure, the absence of V-raising results in the verb remaining in situ and thus following the adverb, as illustrated in (2a). Conversely, when V-raising occurs, the verb moves to T and surfaces before the adverb, as seen in (2b).

(1) a. John often kisses Mary.

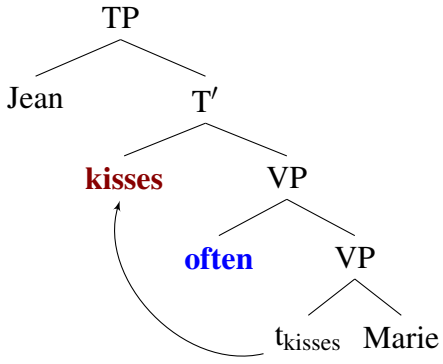
b. *Jean {*souvent embrasse / embrasse souvent} Marie.*
Jean **often** **kisses** / **kisses** **often** Marie
'Jean **often kisses** Marie.'

(Pollock 1989: 367)

(2) a.



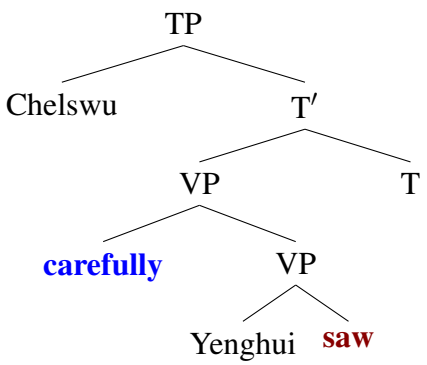
b.



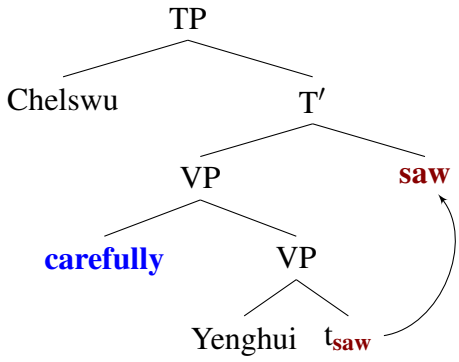
The situation is quite different in head-final languages. In these languages, the relative order between an adverb and a verb does not provide information about V-raising. To illustrate, consider the Korean example in (3). As a head-final language, Korean exhibits a sentence structure in which the verb consistently follows the adverb in non-scrambled sentences. Regardless of whether V-raising occurs, the linear order between the adverb and the verb remains unchanged, since head movement proceeds in a rightward direction, as shown in (4a) and (4b).

(3) *Chelswu-ka **caseyhui** Yenghui-lul **po-ass-ta**.*
C.-NOM **carefully** Y.-ACC **see-PST-DECL**
'C. saw Y carefully.'

(4) a.



b.



As a head-final language, Mongolian exhibits the same property, as shown in (**??**). This sentence follows the canonical SOV word order, confirming its head-final status. Consistent with the Korean structure in (4b), the linear order between the adverb and the verb remains unchanged, since the adverb is adjoined to the left of the VP and any potential head movement would proceed rightward.

(5) *Bat hurdan alim id-sen.*
B. quickly apple eat-PST
Bat quickly ate an apple/apples.

In head-initial languages, the linear order between negation and the verb is also taken as evidence for V-raising. Consider the Icelandic and Danish examples in (6). Koeneman and Zeijlstra (2014) point out that the placement of *sæi* 'saw' to the left of the negation *ekki* in Icelandic (9a) indicates that the verb has undergone V-to-I movement. In contrast, placing the finite verb *fik* 'saw'

to the left of the negation *ikke* leads to ungrammaticality in Danish, as shown in (9b). Thus, there is no evidence for V-to-I movement in Danish.

(6) a. *Ég spurði hvort Jón sæi ekki myndina.*
I asked if Jón saw not the.movie
'I asked if Jón didn't see the movie.' (Icelandic)

b. * *Gad vide om John fik ikke set filmen.*
I wonder if John saw not the movie
'I wonder if John did not see the movie.' (Danish)

(Adopted from Koeneman and Zeijlstra 2014)

Again, the same diagnostic does not work in head-final languages. If a language employs affixal negation, it is impossible to determine the relative order between the verb and negation, as illustrated with Japanese in (7a). Even when a language uses a negation marker that is separate from the verb, the situation does not improve. For example, Mangghuer, a Mongolic language distantly related to Khalkha Mongolian, has a preverbal negation marker *lai*, as shown in (7b). In this case, the configurational relationship between the verb and negation is structurally identical to that between the verb and the adverb in (3). As a result, the verb surfaces to the right of negation regardless of whether V-raising has occurred.

(7) a. *Kanako-wa asagohan-wo tabe-nai.*
K.-TOP breakfast-ACC eat-NEG
'K. does not eat breakfast.' (Japanese)

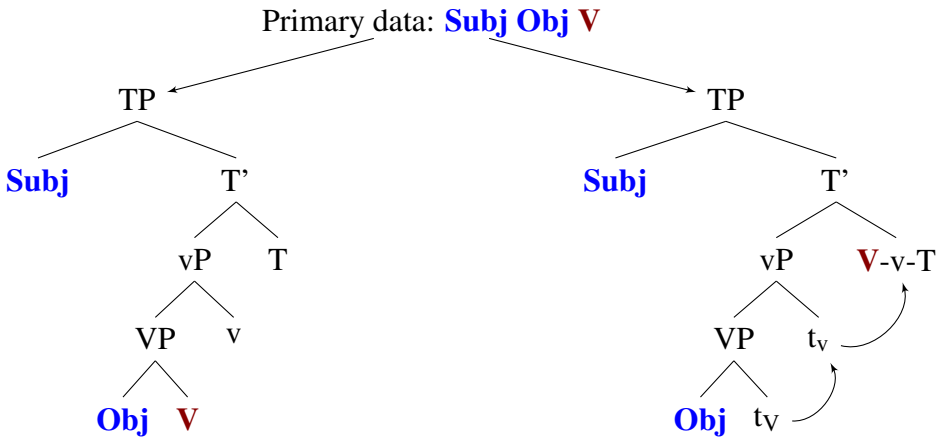
b. *Qi wuge lai maidie-lang.*
you word NEG know-IMPF
'You do not understand language.' (Mangghuer; Adopted from Slater 2005: 144)

As we have seen in this subsection, linear word order does not reveal properties of verb positioning in head-final languages. This raises an interesting question regarding how speakers of head-final languages acquire such grammatical structures.

2.2 Acquisition of Verb Placement Properties in Head-final languages

Han et al. (2007) point out that the string-vacuous nature of verb raising in head-final languages may lead to heterogeneity within a speaker community. To understand this, consider a hypothetical head-final language, as illustrated in (8). When a child acquires the language, they are exposed to primary data exhibiting SOV word order. During the acquisition process, they must determine where verbs are positioned within the grammar they are internalizing. Unlike learners of head-initial languages—where verb placement can be readily inferred from linear word order—learners of head-final languages are not provided with such evidence. As a result, whether or not V-raising occurs may be determined by chance, since both grammars—with and without V-raising—yield the same surface SOV word order, as shown in (8). Drawing on this insight, Han et al. (2007) compellingly argue that Korean exhibits two speaker populations with respect to V-raising: one in which V-raising occurs, and one in which it does not.

(8)



Grammar 1 Grammar 2

However, a split in the speaker population is not the only possible outcome for head-final languages. This is because linear word order is not the only kind of evidence that children may rely on when acquiring V-raising properties.

First, scope interpretation may provide valuable clues about verb placement. It is well known that many head-final languages exhibit rigid surface scope interpretation for quantified phrases (QPs). In these languages, sentences with canonical word order yield only one scope interpretation, derived directly from the surface structure. For now, let us assume that a QP must be c-commanded by another operator in order to fall within its scope (Klima 1964; Miyagawa 2001; Han et al. 2007; Scontras et al. 2017). This assumption implies that the only available scope interpretation for an SOV sentence reflects its surface syntax. Consequently, children may use scope interpretation as indirect evidence for acquiring V-raising properties. As a head-final language, Mongolian is also known for its rigid scope interpretation (Bao et al. 2015; Guntsetseg 2016; Peters 2020), and thus may not exhibit the kind of split population observed in Korean.

Second, a number of studies on subject–verb agreement have proposed that rich agreement morphology may trigger verb raising (Rich Agreement Hypothesis; Bobaljik and Thráinsson 1998; Koeneman and Zeijlstra 2014, among others). Highlighting a strong correlation between rich agreement and V-to-I movement, the theory argues that verbs raise in languages with rich agreement morphology because the relevant agreement features reside in the functional domain and must be checked via verb movement. For example, Koeneman and Zeijlstra (2014) define rich agreement in terms of the presence of binary features for speaker, plurality, and participant in the agreement morphology, and argue that V-to-I movement is triggered by such morphological complexity. Accordingly, the difference between Icelandic and Danish, illustrated in (9) (repeated from (6)), can be explained by the presence or absence of rich agreement: Icelandic exhibits V-to-I movement due to its rich agreement morphology, whereas Danish does not, as shown in (10b).

(9) a. *Ég spurði hvort Jón sæi ekki myndina.*
I asked if Jón saw not the.movie
'I asked if Jón didn't see the movie.' (Icelandic)

b. * *Gad vide om John fik ikke set filmen.*
I wonder if John saw not the movie
'I wonder if John did not see the movie.' (Danish)

(Adopted from Koeneman and Zeijlstra 2014)

(10) a. Icelandic ‘to say’

	Singular	Plural
1	*seg-i*	*seg-jum*
2	*seg-ir*	*seg-ið*
1	*seg-ja*	*seg-ja*

-i → [+speaker], [-plural]
-ir → [-speaker], [-plural]
-jum → [+speaker], [+plural]
-ið → [-speaker], [+participant], [+plural]
-já → [-participant], [+plural]

b. Danish ‘to throw’

	Singular	Plural
1	*kast-er*	*kast-er*
2	*kast-er*	*kast-er*
3	*kast-er*	*kast-er*

-er → [+finite]

Koeneman and Zeijlstra (2014)

If the RAH is correct, it predicts that verbs will remain in situ in Khalkha Mongolian. As shown in (11), Khalkha Mongolian verbs do not inflect for person or number. All the verb forms remain the same—*har-na*—regardless of the person or number of the subject. Thus, it is difficult to even describe a system of agreement morphology in the language, and Mongolian must be considered a language with weak agreement under any version of the RAH. As a result, V-raising is not predicted.

(11) Khalkha Mongolian durative form of *har-* ‘to look’

	Singular	Plural
1	*har-na*	*har-na*
2	*har-na*	*har-na*
3	*har-na*	*har-na*

This theory also leads to an interesting further prediction: if varieties of the same language differ in terms of rich agreement, they should also minimally differ in verb raising (Koeneman and Zeijlstra 2014). Several varieties closely related to Khalkha Mongolian are mutually intelligible but exhibit richer subject–verb agreement morphology. For instance, Khamnigan Mongolian, spoken in Eastern Inner Mongolia, has a rich agreement system, as shown in (12). Thus, if the RAH is correct, Khalkha Mongolian and Khamnigan Mongolian may differ in their V-raising behavior.

(12) Khamnigan Mongolian durative form of *kara-* ‘to watch’ (Janhunen 2006)

	Singular	Plural
1	*kara-nam-bi*	*kara-nam-bide*
2	*kara-nan-ci*	*kara-nan-ta*
3	*kara-nan*	*kara-na-d*

(13)
-bi → [+speaker], [-plural]
-ci → [-speaker], [+participant], [-plural]
-∅ → [-plural], [-participant]
-bide → [+speaker], [+plural]
-tA → [-speaker], [+participant], [+plural]
-d → [-participant], [+plural]

So far, we have seen that learners of head-final languages may rely on indirect cues—such as scope interpretation or rich agreement morphology—in acquiring verb movement properties.

From the researcher's point of view, however, scope interpretation remains the only viable method for investigating V-raising in head-final languages. The reason is straightforward: possible V-raising only becomes visible through changes in linear order, which is not observable in head-final structures. Therefore, in this study, I will use scope interpretation as the primary tool to investigate V-raising in Mongolian.

2.3 Mongolian Properties Regarding Scope Interpretation and V-raising

To determine the syntactic position of verbs in Mongolian through the lens of scope interpretation, it is crucial to first understand the core grammatical properties of the language. In this section, I survey Khalkha Mongolian properties that bear on the applicability of scope interpretation to verb placement, as proposed in Han et al. (2007): frozen scope, object raising, and negation as a clitic. I argue that Khalkha Mongolian satisfies all three.

First, Khalkha Mongolian is known to exhibit rigid scope properties, similar to Korean and Japanese (Guntsetseg 2016; Peters 2020). In these languages, quantified arguments generally receive fixed scope interpretations in canonical SOV word order. Inverse scope becomes available only through scrambling, suggesting that scope interpretation is determined by surface syntactic structure[1]. Accordingly, the availability (or unavailability) of particular scope interpretations should reflect underlying structure.

Second, accusative-marked objects raise above their base-generated position. In Mongolian, objects can appear either as bare NPs or with accusative case marking. Accusative-marked objects are structurally higher than bare NPs, occupying a position above the Voice/vP domain, unless they are heavily stressed or focused (Peters 2020). This explains why a bare NP object appearing to the left of the manner adverb *hurdan* 'quickly' results in unacceptability, as shown in (14a), whereas it is acceptable to the right of the adverb, as in (14b). I therefore treat accusative objects as raised to a functional projection above vP, while bare objects remain vP-internally.

(14) a. *Urnaa alim-iig hurdan id-sen.*
U. apple-ACC quickly eat-PST
'Urnaa ate the apple quickly.'

b. * *Urnaa alim hurdan id-sen.*
U. apple quickly eat-PST
(intended) 'Urnaa ate an apple/apples quickly.'

Third, for negation to serve as a probe for verb position, it must behave as a clitic—base-generated at the VP level and subsequently raised with the verb during head movement (Cinque 1999; Han et al. 2007).

Khalkha Mongolian has several sentential negators: *"ugui* (*güi*), *"ul*, *es*, *b"u"u*, *bitgii*, and *bish*. Among these, *ügüi* and *bish* are postverbal, while the rest are preverbal. *ügüi* is the most productive sentential negator and is used with finite verbal endings derived from participial forms. When *id-sen* 'ate' is negated, it takes the postverbal form *güi*, a reduced, cliticized version of *ügüi*, as shown

[1]Miyagawa (2001) argues that the 'inverse' scope reading available for scrambled sentences is actually a surface scope interpretation of a different syntactic structure. Since the present study does not discuss any scrambled sentences, I will leave this issue in the Khalkha Mongolian context as an open question.

in (15b)[2]. Postverbal negators *b"u"u* and *bitgii* are used only in imperative and optative contexts and are excluded from the present study.

(15) a. *id-sen*
eat-PST
'ate'

b. *id-ee-güi*
eat-PST-NEG
'did not eat'

The preverbal negators *"ul* and *es* occur with finite endings that are incompatible with postverbal *ügüi*. These negators were also excluded from this study because they are obsolete in colloquial usage (Binnick 2011). For example, to negate *ide-v* 'ate' with the canonical past ending *-v*, one must use the preverbal negator *es*, as shown in (16b); using the postverbal *ügüi* is unacceptable, as in (16c). In colloquial speech, however, the negated form with the preverbal negator is never used. Instead, the participial form with the postverbal *güi* in (15b) is used to negate (16a) (Lubsangdorji and Vacek 2004; Kullmann and Tserenpil 2008; Binnick 2011; Gong 2022).

(16) a. *ide-v*
eat-PST
'ate'

b. *es ide-v*
NEG eat-PST
'did not eat'

c. * *ide-v-güi*
eat-PST-NEG
(intended) 'did not eat'

I further assume that the postverbal negator *bish* occupies its own projection above the TP domain and is therefore excluded from the scope of this study. Bao et al. (2015) argues that *bish* always takes wide scope over universal quantifiers, regardless of the position of the quantified argument. This implies that *bish* occupies a syntactically high position capable of c-commanding the subject. Additional evidence comes from the clitic status of *ügüi*. I will return to this issue shortly, but first, I explain why the clitic nature of negation matters in this context.

Cinque (1999) argues that cliticized negation in some Romance languages is base-generated in a specifier position and "rides on" the verb through head movement, targeting the landing site of verb raising. Han et al. (2007) extend this account to short negation in Korean, which is adjoined at the VP level and undergoes movement with a complex verbal head. Consequently, the negation reaches the landing site of verb raising, and its scope is determined from that position in combination with the frozen scope property.

Why should we consider *ügüi* a clitic? First, it can be separated from the verb it negates by a focus clitic *c*, as shown in (17). According to the affix–clitic distinction in Pullum and Zwicky

[2]Some readers may notice that the past suffix has changed to *-ee*. Although it is possible to attach *güi* to the original past suffix *-sen*, doing so introduces a presupposition that the act of not eating contradicts the speaker's expectation (Binnick 2011). I therefore treat *-ee-güi* as the negative counterpart of *-sen*, as seen in the relation between (15a) and (15b).

1983, an affix cannot appear after a clitic. Since the focus marker *c* is unambiguously a clitic, this separation suggests that *ügüi* is also a clitic.

(17) *Camaas öör hün-iig hairla-j cada-h* ***c*** ***ügüi****, ganchan minii boddog yum.*
you.ABL another person-ACC love-CVB can-NPST FOC NEG only my thought PTL
'My only thought is that I can't fall in love with anyone else apart from you.' (adopted from Dagvasumberel 2015)

Second, the separation of *ügüi* from the verb does not trigger insertion of a light verb, as would be expected if the negator were affixal. In Japanese, when negation is separated from the verb stem by an emphatic particle, a light verb *suru* 'do' is inserted to host the negation, as shown in (18) (Kuroda 1965; Miyagawa 2001). The insertion of the light verb reflects the affixal status of negation. In contrast, no such insertion occurs with *ügüi*, as seen in (17).

(18) *Sensei-o zen'in-ga seme-****mo*** ***si-nakat****-ta.*
tacher-ACC all-NOM blame-even do-NEG-PST
'The teacher, all did not even blame.' (adopted from Miyagawa 2001)

Third, the availability of double negation in Khalkha Mongolian supports the claim that *ügüi* is an adverbial clitic, rather than the head of a NegP projection. This property was previously used to argue for the adverbial nature of short negation in Korean (Han and Lee 2007; Han et al. 2007; Choi and Harley 2019). The co-occurrence of *ügüi* and *bish* yields a double negation interpretation, as shown in (19).

(19) *Bat alim id-ee-güi bol bish.*
B. apple eat-PST-NEG TOP NEG
'It is not the case that Bat didn't eat an apple/apples' or 'Bat didn't NOT eat an apple/apples.'

Finally, negative suppletion involving *bai-* 'to be, to exist' shows that negation must originate in a position structurally close to the verb root. Mongolian verbs generally do not exhibit suppletive alternations. However, *ügüi-sen* 'were not' may be used in place of the regular negative form *bai-gaa-güi*, as shown in (20b). Furthermore, it is not possible to further negate (20b) with postverbal *güi*, as shown in (20c). I interpret this as evidence that the negator is close enough to the stem to trigger suppletion. In (20a), the past tense morpheme *-gaa* intervenes between *bai-* and the negation, suggesting that structural adjacency is required to condition root allomorphy.

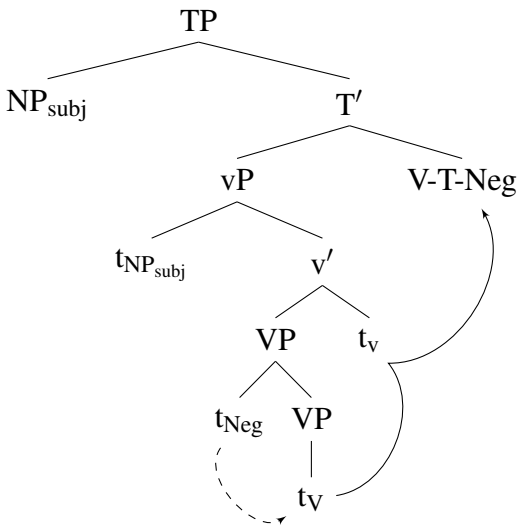
(20) a. *Chi bai-gaa-güi bol ...*
you be-PST-NEG COND ...
'If you were not (there) ...'

b. *Chi ügüi-sen bol ...*
you be.not-PST COND ...
'If you were not (there) ...'

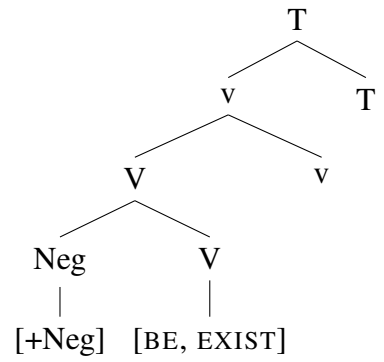
c. *Chi ügüi-sen-güi bol ...*
you be.not-PST-NEG COND ...
(intended) 'If it was not the case that you were not (there) ...'

Accordingly, I assume that *ügüi* is base-generated vP-internally and is linearly dislocated to the left edge of the verb post-syntactically, as shown in (21). In the case of negative suppletion, I follow Chung (2009) in assuming that the Neg node fuses with the V node when its sister is the verb BE, EXIST, as shown in (22). This explains why the additional negator *güi* is ungrammatical in (20c): the negation has already been exponed, and no further Vocabulary Insertion is possible.

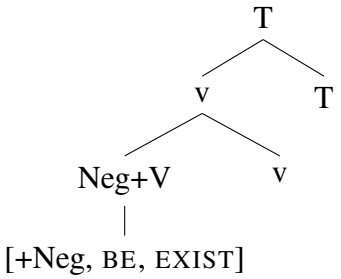
(21)



(22) a.



b.



For the purposes of this study, I assume that the negator *ügüi* is a clitic. This assumption is supported by several empirical observations: it can be separated from the verb by a focus clitic; it does not trigger light verb insertion; it co-occurs with another negator in double negation; and it participates in root suppletion, indicating structural locality to the verb. Taken together, these facts show that *ügüi* is not the head of NegP but a clitic adjoined at the vP level, making it a suitable diagnostic for verb position in Mongolian.

In addition to the three criteria proposed by Han et al. (2007), I further assume that the subject moves to Spec,TP to satisfy the EPP (Gong 2022; Lim 2023). In Khalkha Mongolian, subjects cannot appear to the right of low adverbials such as manner or frequency adverbs, as shown in (23) (Guntsetseg 2016).

(23) a. *(*Hurdan) Bilgüün (hurdan) ene baishin-g (hurdan) bar'-san.*
quickly B. quickly this building-ACC quickly build-PST
'Bilguun built this building quickly.'

b. *(*Dandaa) Tuyaa (dandaa) Dorj-iig (dandaa) tusal-dag.*
always T. always D.-DAT always hel-HAB
'Tuyaa always helps Dorj.'

The linear order between subjects and higher adverbs like temporal adverbs reveals that it is acceptable to place the subjects to the right of the adverbs, as shown in (24) (Lim 2023). Putting it altogether, the distribution shown in (23) and (24) shows that the position of subjects is Spec,TP.

(24) *Margaash Naraa ene baishin-g nuraa-n.*
tomorrow N. this building-ACC tear.down-NPST
'Tomorrow N. will tear down this building' (Lim 2023)

2.4 Conflicting claims on scope interpretations

Given that Khalkha Mongolian satisfies the key prerequisites for using scope interpretation as a diagnostic—namely frozen scope, cliticized negation, and object raising—we now turn to how scope judgments can be used to probe verb position. To clarify the logic behind the methodology before applying it to Mongolian, we begin with an example in English, where scope interpretations between negation and a universally quantified phrase give rise to distinct truth conditions.

Let us consider the English sentence *Every fox didn't chase the marmot*, a negative sentence with a subject QP. This sentence is known to have two possible interpretations: the *neg* $>$ $\forall$ (surface) scope reading and the $\forall >$ *neg* (inverse) scope reading. Suppose there are three foxes and one marmot in the given context.

First, the $\forall >$ *neg* reading follows directly from the syntactic structure: the universal quantifier takes scope over negation ($\forall >$ Negation), as the subject QP c-commands the negation. Under this reading, the sentence is TRUE when none of the three foxes chased the marmot. This is commonly referred to as the "surface reading" because it reflects the surface structure.

The *neg* $>$ $\forall$ reading, by contrast, involves an inverse scope configuration, with negation taking scope over the universal quantifier. Since this interpretation does not correspond to the surface structure, it is referred to as the "inverse reading." In this case, the sentence is TRUE when some but not all of the three foxes chased the marmot.

However, using scope interpretation to detect V-raising is not without challenges: native speakers' scope judgments can be inconsistent. While some languages, such as Mandarin Chinese, allow only surface scope interpretations (Huang 1981; Huang 1982; Lee 1986; Aoun and Li 1989; 2003; pace Zhou and Gao 2009), many others permit both readings. But this does not mean that they are equally accessible. Speakers may show a strong preference for one interpretation over the other (e.g., Scontras et al. 2017 for English), and some may fail to entertain the less preferred reading if contextual cues are weak. In other cases, as seen in Korean (Han and Lee 2007), the speaker population may even split into distinct groups based on scope judgments. Therefore, it is risky to draw conclusions about a language's scope properties based on judgments from just a few speakers.

This difficulty is reflected in the conflicting claims found in prior literature. Han et al. (2007) argue that it is impossible to reach a convergent conclusion from the scope judgments reported in previous work—an observation that motivated their experimental methodology.

Similar inconsistencies are found in the Mongolian literature. Guntsetseg (2012) reports that the universal quantifier *b"uh* exhibits scope ambiguity with the negator *ügüi*, though her discussion focuses only on subject QPs. Thus, it is unclear whether the same holds for object QPs. Bao

et al. (2015) argue that scope is determined by the type of negation marker. Focusing on *"ugei*[3] is an alternative spelling of *ugui*, primarily used in Inner Mongolia, where Mongolian is written in traditional script. Since Bao et al. (2015) analyze the Khorchin dialect, their transliteration differs from the Cyrillic-based spelling used here. The difference between *"ugei* and *ügüi* is thus purely orthographic. and *bisi*, they claim that the former permits only the $\forall > neg$ reading while the latter permits the $neg > \forall$ reading, regardless of whether the quantified argument is a subject or object. Contradicting this, Gong (2022) reports that the $neg > \forall$ reading is possible for universally quantified objects, though she concurs with Bao et al. (2015) on the judgments for subject QPs. Finally, (Guntsetseg 2016: 96) suggests that scope interpretation may not always align with syntactic structure: accusative-marked indefinite objects are interpreted higher than linearly preceding frequency adverbs. Table 1 summarizes the scope judgments reported in the literature.

	Guntsetseg (2012)		Bao et al. (2015)		Gong (2022)	
	$\forall > neg$	$neg > \forall$	$\forall > neg$	$neg > \forall$	$\forall > neg$	$neg > \forall$
Subject	yes	yes	yes	no	yes	no
Object	No data	No data	yes	no	?	yes

Table 1: Mongolian scope judgments reported in the previous literature

Inconsistent judgments were also observed among native speakers. As a pilot study for the current investigation, I collected scope judgments from three native speakers for negative sentences with quantified arguments. As shown in Table 2, considerable variation emerged for sentences with universally quantified subjects and negation, while all three speakers judged both readings to be available for universally quantified objects. These findings confirm that the variation observed in the literature is genuine. Importantly, none of the individual consultants replicated the exact pattern reported in any single prior study: whenever a subject QP pattern matched one source, the object QP judgments diverged from it.

	Khalkha 1		Khalkha 2		Khorchin 1	
	$\forall > neg$	$neg > \forall$	$\forall > neg$	$neg > \forall$	$\forall > neg$	$neg > \forall$
Subject	yes	yes	yes	no	no	yes
Object	yes	yes	yes	yes	yes	yes

Table 2: Scope judgments elicited in my fieldwork

3 Experimental Investigations

As we have seen, scope judgments are difficult to systematically elicit—even though, in principle, they provide a valuable diagnostic for investigating V-raising. As Han et al. (2007) point out, discrepancies in reported judgments may stem from "insufficient discourse context." That is, if the

[3] *"u*gei

interpretive context is tightly controlled, more consistent judgments may be obtained from native speakers. For this reason, I adopt an experimental approach in the present study.

3.1 Methods

3.1.1 The Truth-Value Judgment task

This study employed a truth-value judgment task (TVJT) (Crain and Thornton 1998) to elicit scope judgments from native speakers of Mongolian. Originally developed to investigate children's grammatical competence, this method is known for minimizing metalinguistic interference that can obscure a speaker's underlying intuitions.

Scope judgments were collected as follows. In each test trial, participants were presented with a sentence and an accompanying picture. The sentence was a negative sentence containing either a universally quantified subject NP (subject QP) or object NP (object QP). The picture illustrated one of the two possible scope readings available for the sentence. Participants were then asked to judge whether the sentence was true in the context represented by the picture. A "true" response indicates that the scope reading encoded in the picture is available in the participant's grammar. Conversely, a "false" response suggests that the depicted reading is not available for the given sentence.

The reliability of this method is grounded in an important assumption about how speakers evaluate ambiguous sentences in controlled contexts. This method relies on the assumption that participants will give the speaker of a sentence the benefit of the doubt (Lidz and Musolino 2002; Han et al. 2007). That is, regardless of how dispreferred a particular reading may be, participants will judge the sentence as TRUE if that reading is available in context. For example, it is well known that English speakers tend to prefer surface scope interpretations over inverse scope interpretations in doubly quantified sentences (Scontras et al. 2017). In such cases, a speaker might reject the inverse reading when presented with the sentence in isolation, due to the lack of contextual support. However, when an appropriate context is provided—such as through pictures, as in the present study—the same speaker may accept the reading, even if it is only weakly accessible. In other words, participants will judge a sentence FALSE only when the scope reading represented in the picture is truly unavailable in their grammar.

3.1.2 Participants

I tested 60 adult speakers of Khalkha Mongolian (aged 18–50, M = 25.82; fifty-three females and seven males).

3.1.3 Experimental Design and materials

Factors

Two factors were tested, each with two levels: scope interpretation (`Scope`; *neg* $>$ $\forall$ vs. $\forall >$ *neg*) and the type of quantified argument (`QP`; subject QP vs. object QP). The `Scope` factor was treated as a between-participants factor in order to minimize potential interference from a more salient or preferred reading, if such a bias exists. That is, each participant was assigned to only one level of `Scope`—either *neg* $>$ $\forall$ or $\forall >$ *neg*—but not both. The `QP` factor was treated as a within-participants factor: each participant was presented with both subject QP and object QP

sentences. The experiment thus consisted of four conditions in total, as summarized in Table 3. Thirty participants were assigned to each of the two levels of the Scope factor.

2×2 design:	Scope ($neg > \forall$ vs. $\forall > neg$)	$\times$	QP (subject vs. object)

Scope	QP	
$neg > \forall$	Subject QP Object QP	$n = 30$
$\forall > neg$	Subject QP Object QP	$n = 30$

Table 3: Experimental design

Procedure

Each participant completed 10 test sentences for the subject QP condition and 10 for the object QP condition, along with 20 filler sentences. All items were presented in pseudo-randomized order. Participants were also given four practice trials to familiarize themselves with the task before the main session began.

For each item, participants were presented with a sentence and a corresponding picture. They were asked to judge whether the sentence was true in the depicted situation. An example trial is illustrated in (1), and the task instructions are provided in (25). After making a decision, participants selected either the "True" or "False" button on the screen. To ensure commitment to each response, participants could not proceed to the next stimulus until they clicked the "Continue" button.

Figure 1: Experiment screen

(25) a. Instruction

Daraah zurg-iig dürsel-sen ögüülber zöv üü?
following picture-ACC depict-PST.PTCP sentence correct Q

'Is the sentence that describes the following picture correct?'

b. Stimulus sentence
Ter böö büh tüüdeg gal asaa-gaa-güi
that shaman every bonfire fire kindle-PST-NEG
'That shaman did not make every bonfire.'

c. Response
Ünen / Hudal
True / False

d. Continue button
Ürgeljlüüleh
continue

Material All experimental sentences were past-tense transitive clauses. I used sentences in the past tense because past-tense constructions provide more accurate information about the syntactic position of negation in Mongolian than present-tense clauses, due to language-specific properties. Since participants were asked to judge the truth value of a sentence in relation to a picture, the depicted situation could be interpreted as either ongoing or past. In Mongolian, the simple present suffix *-dAg* conveys habitual meaning, as shown in (26a). The present progressive meaning is expressed by a combination of a main verb suffixed with the converbial marker *-J* and the auxiliary verb *bai-* 'be' in the non-past tense, as illustrated in (26b).

(26) a. *Bi ih surguul'-d ajilla-dag.*
I big school-DL work-HAB
'I work at the university.' ((Kullmann and Tserenpil 2008: 144))

b. *Bagsh tend suu-j bai-na.*
teacher there sit-CVB be-NPST
'The teacher is sitting there.' ((Kullmann and Tserenpil 2008: 136))

The issue is that negation in the present progressive is not attached to the main verb but to the auxiliary *bai-*, as in (27). Furthermore, the main verb and auxiliary are often separable; for example, the particle *l* 'only' can intervene between them. Since the negation is suffixed to the auxiliary, it fails to inform us about the position of the main verb.

(27) *Bagsh tend suu-j bai-h-güi.*
teacher there sit-CVB be-NPST-NEG
'The teacher is not sitting there.'

n contrast, past-tense constructions are morphosyntactically simpler: the negation is cliticized to the main verb itself, as shown in (28). In this case, negation moves together with the complex verbal head. If V-raising occurs in Mongolian, this configuration permits the complex head to raise without structural restriction. Therefore, as long as participants interpret the visual scenario as past, their judgments offer meaningful information about the syntactic position of the verb-negation complex.

(28) *Bi 2023 on-d Mongol-d ir-ee-güi*
I 2023 year-DL Mongolia-DL go-PST-NEG
'I did not go to Mongolia in 2023.'

I do not elaborate further on how negation attaches to auxiliary verbs, but as discussed in Section 2, I assume independently that negation is adjoined to and cliticized on the main verb. Since the structure of present progressive clauses lies beyond the scope of this study, I restricted all stimuli to past-tense constructions, which have a simpler form and are more compatible with the current research goals.

Twenty items, each consisting of two nouns and one verb, were used to construct the stimulus sentences. These elements were inserted into five sentence templates—two for target sentences and three for fillers. All templates were simple transitive sentences. I detail the structure of each sentence template below.

Based on these sentences, images were generated using ChatGPT (OpenAI 2024). Each image was prompted to visually represent the meaning of the corresponding sentence in a simple 16:9 grayscale format. Because ChatGPT cannot reliably produce sequences representing events over time, the items were designed to describe bounded events that could be depicted in a single frame. Given the temporal reference of the test sentences, participants were expected to interpret each image as depicting a past situation. The availability of unbounded readings for the simple past varies cross-linguistically and across verb classes. For instance, (Folli and Harley 2005: 111), citing Giorgi and Pianesi (1997), notes that Italian creation and consumption verbs may allow unbounded event interpretations. In such languages, a sentence like *John ate an apple* might be judged true even when the picture shows a man still eating the apple. However, in languages that do not allow such readings, the same sentence would be judged false in such a scenario. It remains unclear whether the simple past in Mongolian permits unbounded readings; this question lies beyond the scope of this study. By contrast, the bounded reading of the past is cross-linguistically robust (Velupillai 2012: 211).

To prevent participants from basing their truth-value judgments on perfectivity rather than scope, I designed all items so that a bounded event reading could be easily supported by the visual stimuli. Consider (29). It is straightforward to visualize sentence (29a) as a bounded event: a child sits in front of multiple apples and apple cores. Because eating apples leaves a visually salient result (unless the person eats even the core), such an image clearly supports boundedness. In contrast, sentence (29b) is harder to visualize in this way. If the child ate several dumplings, the aftermath may not be visually obvious. Depicting half-eaten dumplings might imply an ongoing event, prompting participants to judge the sentence false—not due to scope, but because of an apparent mismatch with past tense.[4].

(29) a. Triplet: {*hüühed* 'child,' *alim* 'apple,' *id-* 'to eat'}
Ter hüühed büh alim-iig id-ee-güi
that child every apple-ACC eat-PST-NEG
'That child did not eat every apple.'

b. Triplet: {*hüühed* 'child,' *buuz* 'dumpling,' *id-* 'to eat'}
Ter hüühed büh buuz-iig id-ee-güi
that child every dumpling-ACC eat-PST-NEG

[4]A researcher might consider adding cues like a child's facial expression of fullness. However, one of the motivations behind this experimental method is to minimize variation in participants' ability to infer contextually appropriate scope readings. Relying on such indirect visual signals undermines this goal: how can we be sure that all participants will interpret a child's facial expression as indicating satiety?

'That child did not eat every dumpling.'

The comprehensive list of items used in the experiment is provided in Appendix I.

The visual stimuli were designed to ensure that truth-value judgments reflected scope interpretation alone, not pragmatic confounds such as relevance failure or infelicity (**?**Heim 1983). For a sentence involving negation and a universally quantified object to be meaningfully evaluated with respect to scope, the relevant event domain must be explicitly present in the image.

Take sentence (29a) as an example. To evaluate its scope reading, the scene must depict all relevant semantic components—a child, multiple apples, and an eating event. If no eating occurs (e.g., the child is asleep), participants might judge the sentence false due to presupposition failure, not because of scope interpretation.

To avoid such confounds, each image explicitly depicted the full semantic triplet from the corresponding sentence: subject (agent), object (theme), and event. This design ensured that truth-value judgments could only be based on the scope relation between negation and quantification.

Finally, let us look more closely at the design of the experimental sentences by condition. First, the `Subject QP` condition tested whether native speakers allow both *neg* $>$ $\forall$ and $\forall$ $>$ *neg* readings for negative sentences with universally quantified subjects. Items were constructed by inserting each triplet into a template such as (30a). The subject was quantified with *büh* 'every.' The object appeared as a bare NP, since its structural position is irrelevant in this condition. In Mongolian, bare NP objects are typically interpreted as pseudo-incorporated and display number neutrality (Guntsetseg 2012). This flexibility allowed greater latitude when designing the images. An example of a test sentence in this condition is shown in (30b).

(30) a. *Büh* Subj Obj V-AA-*güi*
every Subj Obj V-PST-NEG
'Every subject did not verb an object/objects'

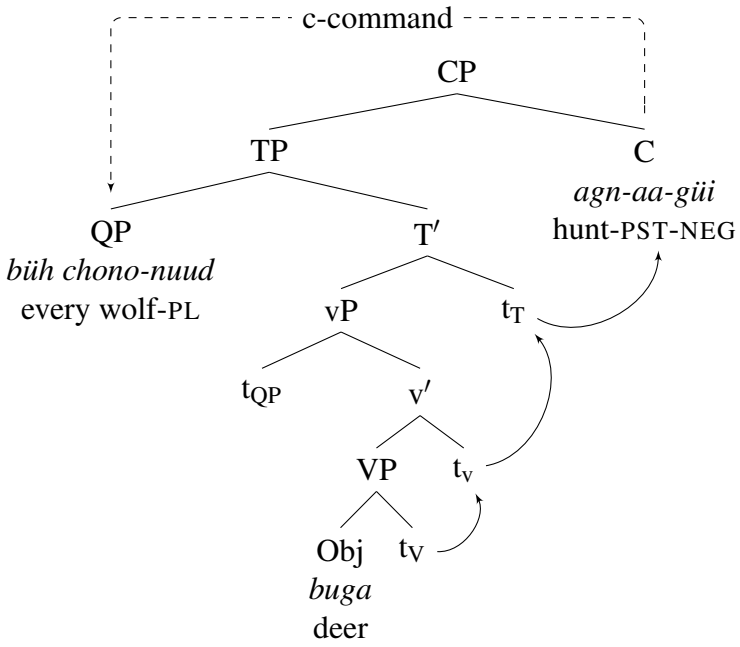
b. *büh* *chono-nuud buga agn-aa-güi*
every wolf-PL deer hunt-PST-NEG
'Every wolf did not hunt a deer/deer.'

Based on the test sentences, participants in each `Scope` group were assigned a picture corresponding to the intended scope reading. Participants in the *neg* $>$ $\forall$ group saw pictures where some but not all referents of the subject performed the action denoted by the verb on the object referents. For instance, with (30b), participants were shown a picture in which some but not all wolves hunted deer (Figure 2). If a participant judged the sentence to be true in that context, I interpreted this as evidence that the *neg* $>$ $\forall$ reading is available for that sentence.

In syntactic terms, this corresponds to a configuration where the negation c-commands the universally quantified subject. Assuming that the subject occupies Spec,TP, this entails that the verb has moved to a position higher than TP—i.e., V-to-C movement—illustrated in (31).

(31)

Figure 2: Picture for `Subject QP` $\times$ *neg* $>$ $\forall$ condition

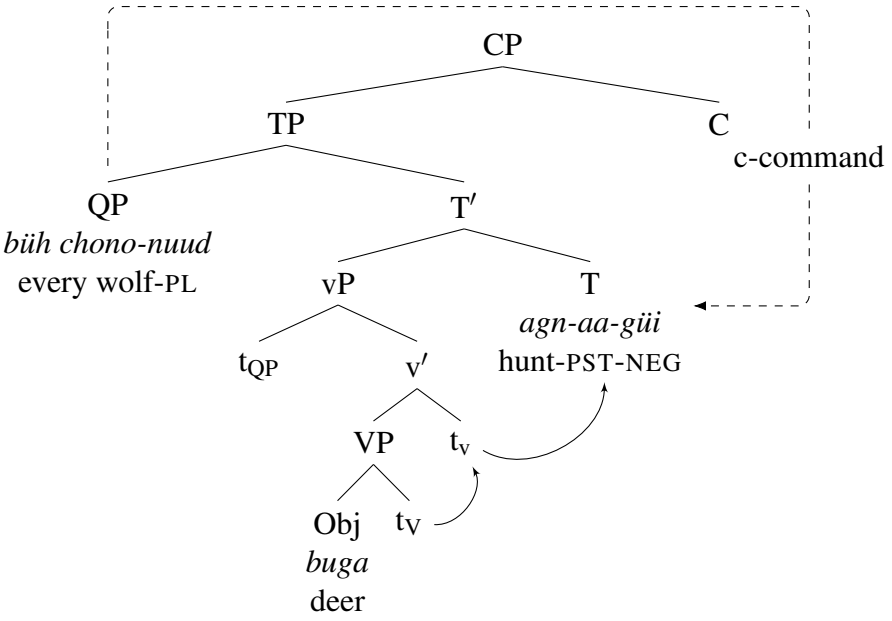


Participants in the $\forall > neg$ group were shown pictures where none of the subject referents carried out the action on any of the object referents. For example, in the case of (30b), the picture depicted a group of wolves not engaging in any hunting activity (Figure 3). As with the *neg* $>$ $\forall$ group, if the participant judged the sentence to be true, I took this to indicate that the $\forall > neg$ reading is available.

Structurally, this implies that the subject QP c-commands the negation. Regarding verb movement, this entails that the verb does not raise higher than TP. Multiple configurations are compatible with this reading: the verb may remain in situ, undergo V-to-T movement, or move to an intermediate functional head—all configurations where the verbal complex remains c-commanded by the subject QP, as illustrated in (32, 33, and 34). However, as will be discussed in Section 4, there exists an alternative derivation for the $\forall > neg$ reading.

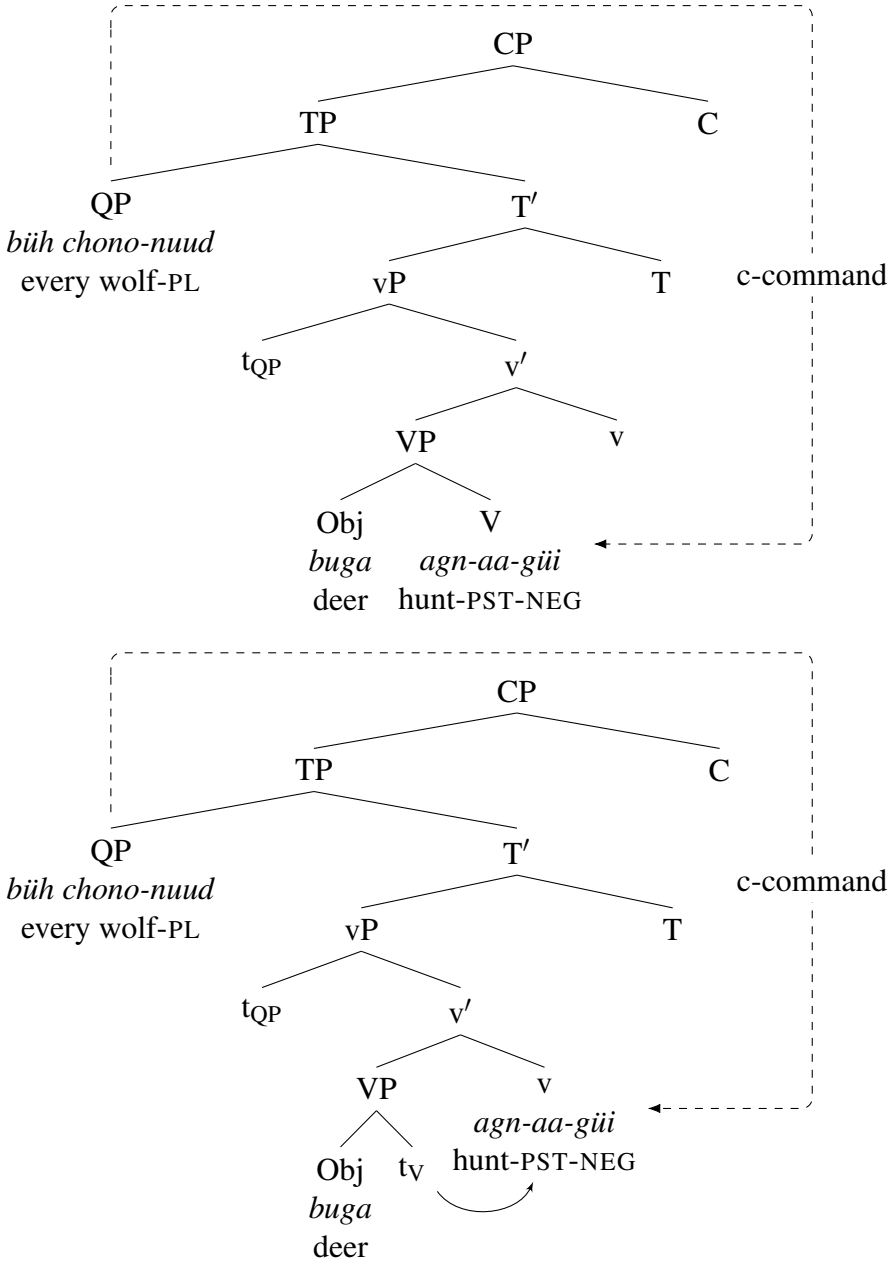
(32)

Figure 3: Picture for `Subject QP` $\times \forall >$ *neg* condition



(33)

(34)



Let us now consider the object QP condition, which tests whether *neg* $>$ $\forall$ and $\forall$ $>$ *neg* readings are available in negative sentences containing universally quantified objects. The structure and an example are shown in (35a) and (35b), respectively. The subject is modified by a distal demonstrative ('that') and interpreted as a singular definite, while the object bears the universal quantifier *büh* and accusative case.

(35) a. *Ter* Subj *büh* Obj-*iig* V-AA-*güi*
that Subj every Obj-ACC V-PST-NEG
'That subject did not verb every object.'

b. *Ter chono büh bug-iig agn-aa-güi*
that wolf every deer-ACC hunt-PST-NEG
'Every wolf did not hunt a deer/deer.'

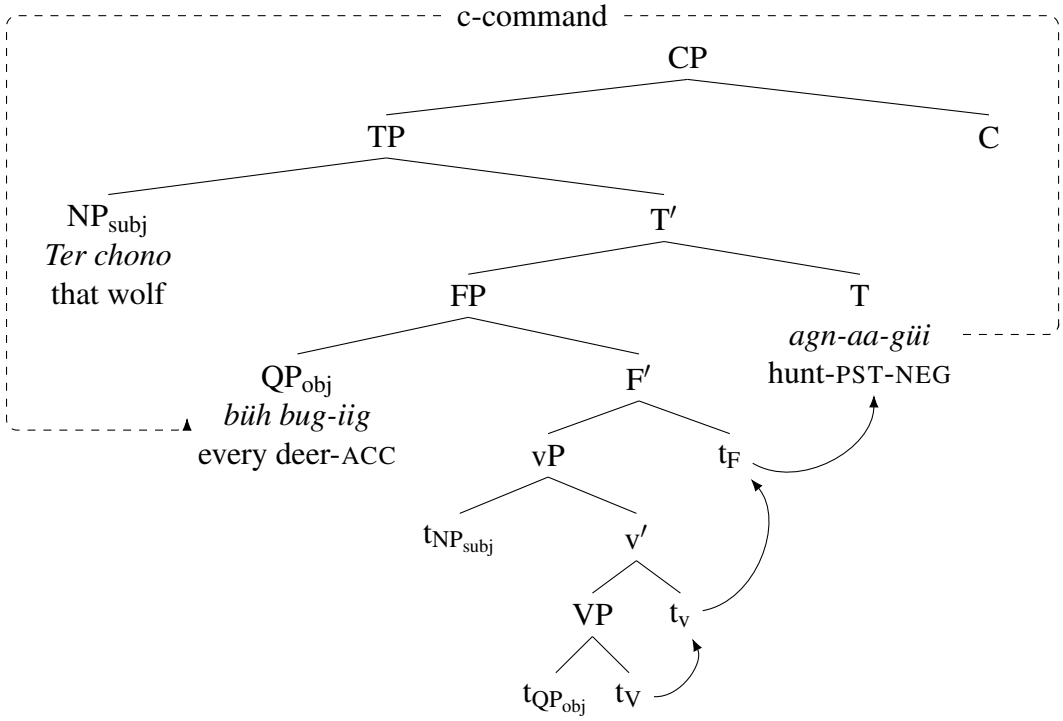
Following Peters (2020), I assume that accusative-marked objects are raised to a functional projection (FP) above the vP domain. In this position, the object is structurally higher than the verb. If V-raising does not occur, the object asymmetrically c-commands the complex verbal head. If the object were to remain in situ, it would stand in a symmetric c-command relation with the verb, potentially allowing both scope readings and making interpretation of verb position more ambiguous. By assuming object raising, we secure an asymmetry in the structure that makes scope-based diagnostic inferences more reliable.

As in the subject QP condition, participants were divided by `Scope` group. Participants in the *neg* $>$ $\forall$ group saw pictures in which the subject referent carried out the action on some but not all of the object referents. For (35b), a wolf hunting some—but not all—deer was depicted (Figure 4). If a participant judged the sentence to be true, I concluded that the *neg* $>$ $\forall$ reading was available.

Figure 4: An example picture of Object QP $\times$ *neg* $>$ $\forall$ condition

In syntactic terms, this implies that the complex verbal head raises above FP and c-commands the object. This entails at least V-to-T movement, as shown in (36).

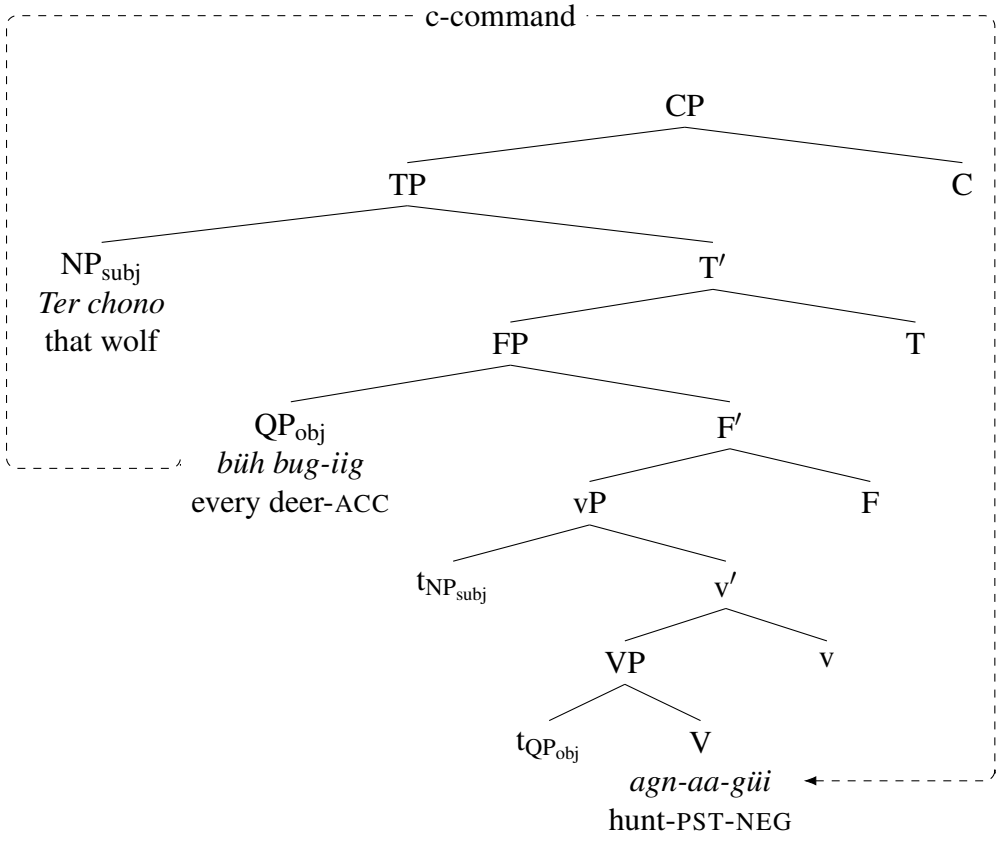
(36)



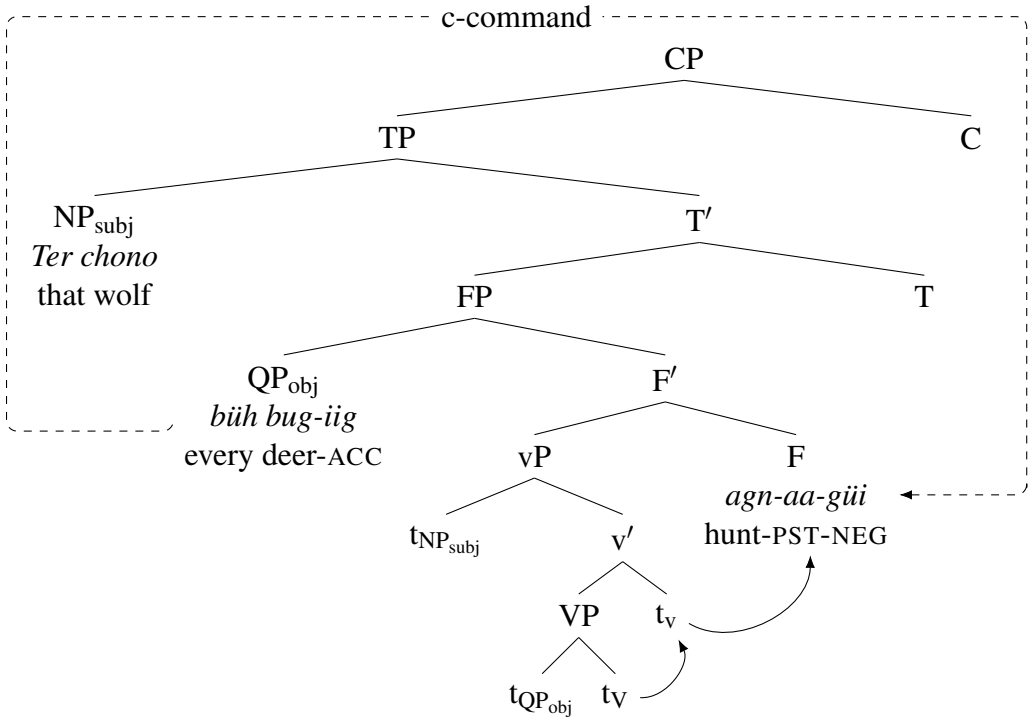
Participants in the $\forall > neg$ group were shown pictures representing the $\forall > neg$ reading, where the subject referent performed the action on none of the object referents. For instance, with (35b), the corresponding picture depicted a wolf and multiple deer, with no interaction. If the participant judged the sentence to be true under this condition, I interpreted this as evidence for the availability of the $\forall > neg$ reading.

This implies that the verbal complex remains lower than the object QP in Spec,FP. Three structural possibilities are compatible with this reading: the verb remains in situ, raises to v, or raises to an intermediate head F, as shown in (37, 39, and 38). In all these cases, the object QP c-commands the verb. As with the subject QP condition, I acknowledge that there is another way to derive the $\forall > neg$ reading, which I address in Section 4.

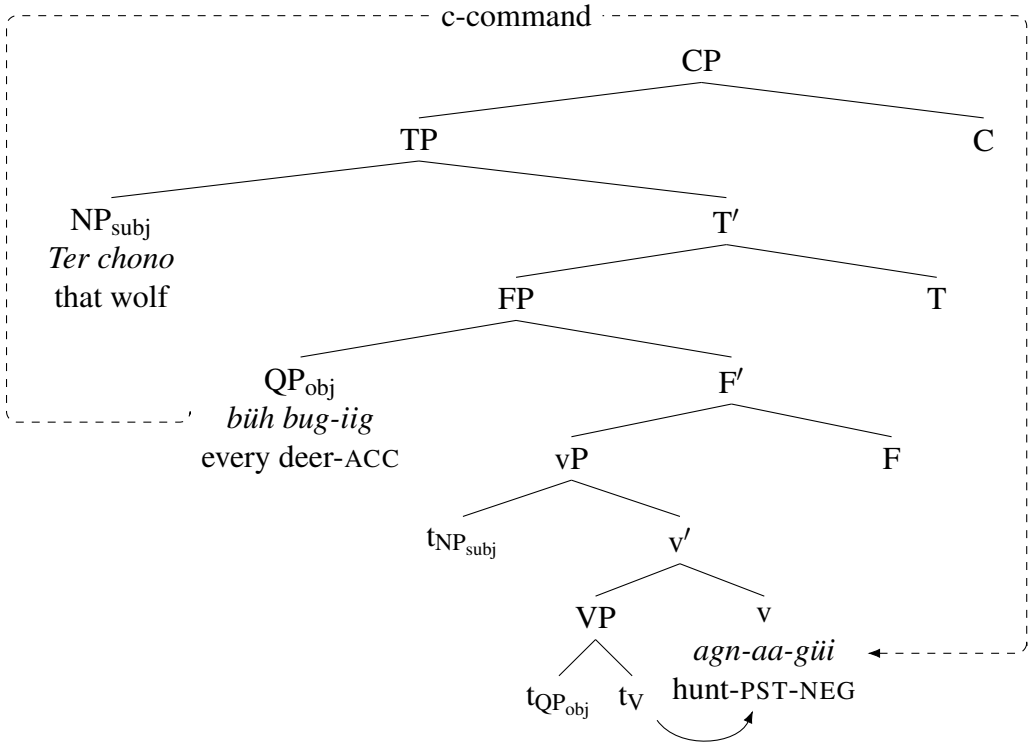
(37)


c-command
CP
TP
C
NP$_{subj}$
Ter chono
that wolf
T′
FP
T
QP$_{obj}$
büh bug-iig
every deer-ACC
F′
vP
F
t$_{NP_{subj}}$
v′
VP
v
t$_{QP_{obj}}$
V
agn-aa-güi
hunt-PST-NEG

(38)


c-command
CP
TP
C
NP$_{subj}$
Ter chono
that wolf
T′
FP
T
QP$_{obj}$
büh bug-iig
every deer-ACC
F′
vP
F
agn-aa-güi
hunt-PST-NEG
t$_{NP_{subj}}$
v′
VP
t$_v$
t$_{QP_{obj}}$
t$_V$

(39)



As mentioned earlier, each participant also completed twenty filler sentences. The fillers served two purposes: to assess the participant's attention and to evaluate their understanding of negation and universal quantification. There were four types: (i) simple affirmative transitive sentences (40a), (ii) simple negative transitive sentences (40b), (iii) affirmative transitive sentences with a subject QP (40c), and (iv) relevance-failure sentences. The fourth type shared the structure of the first but was paired with an unrelated picture, intentionally introducing a relevance failure.

(40) a. *Ter oyutan bömbög shid-sen.*
that student ball throw-PST
'That student threw a ball/balls.'

b. *Ter zaluu hon' hyarg-aa-güi*
that young.person sheep shear-PST-NEG
'That young man did not shear a sheep/sheep.'

c. *büh högsh-chüüd üher hötöl-sön.*
every old-PL cow lead-PST
'Every old person led a cow.'

Each filler sentence was paired with a corresponding image. Among the five fillers from the first three types, three were designed to be judged TRUE and two FALSE.

3.2 Results

Based on participants' responses to the filler sentences, I calculated an accuracy rate for each participant, defined as the proportion of filler items correctly judged. Participants whose accuracy fell more than one standard deviation below the mean (i.e., below 78.0%) were excluded from

analysis. This exclusion criterion was based on the empirical distribution of the full sample (M = 87.9%, SD = 9.9%).

For each test condition, I calculated the proportion of sentences judged TRUE by each participant as the acceptance rate. These data are summarized in Table 4 and visualized in Figure 5. Mean acceptance rates across all conditions exceeded 84.0%.

QP	Scope	Acceptance rate (%)
Subject QP	$neg > \forall$	87.0
	$\forall > neg$	89.3
Object QP	$neg > \forall$	85.3
	$\forall > neg$	84.0

Table 4: Mean acceptance rates by condition

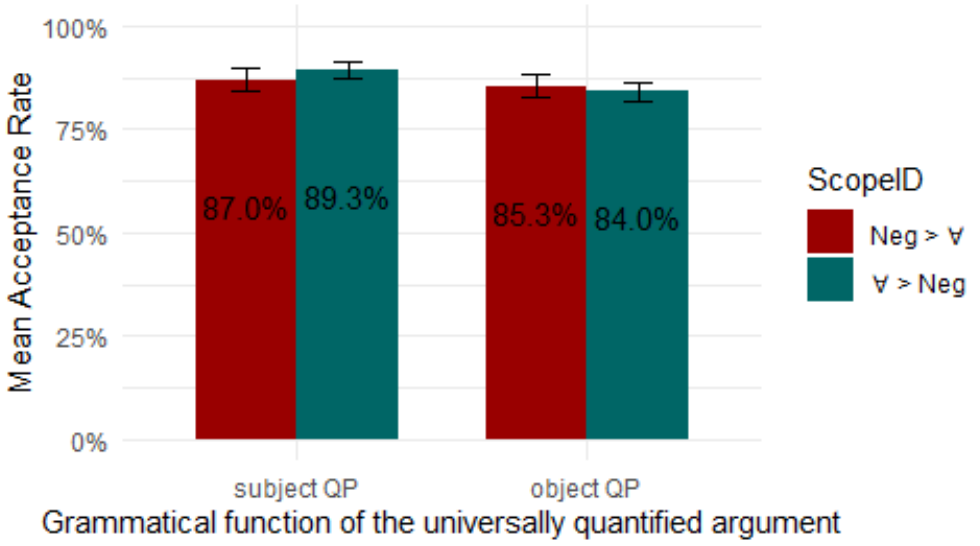


Figure 5: Mean percentage acceptance rates

To evaluate whether each scope interpretation was systematically available, I conducted one-sample t-tests (with Wilcoxon signed-rank tests for robustness) comparing acceptance rates to chance (0.5). The Shapiro–Wilk test revealed significant deviations from normality in all conditions (all $p < .01$).

Despite non-normality, one-sample t-tests showed that mean acceptance rates in all conditions were significantly above chance ($p < .001$). Wilcoxon signed-rank tests confirmed these findings (all $p < .001$), indicating that both $neg > \forall$ and $\forall > neg$ scope interpretations were reliably accessible across configurations. Summary statistics are reported in Table 5.

To test for differences between conditions, I fitted a Generalized Linear Mixed-Effects Model (GLMM) using the `lme4` package in R. The binary outcome (TRUE = accepted) was modeled using a logit link and binomial distribution. Fixed effects included the grammatical function of the quantified phrase (`QP`), the scope interpretation (`Scope`), and their interaction.

Condition	t(df)	p-value	95% CI
$neg > \forall \times$ subject QP	$t(29) = 11.60$	$< .001$	[0.79, 0.90]
$neg > \forall \times$ object QP	$t(29) = 12.50$	$< .001$	[0.80, 0.91]
$\forall > neg \times$ subject QP	$t(29) = 10.95$	$< .001$	[0.78, 0.89]
$\forall > neg \times$ object QP	$t(29) = 11.10$	$< .001$	[0.78, 0.89]

Table 5: Results from one-sample t-tests

Random effects included a random intercept and by-participant random slope for `QP`, as well as random intercepts and random slopes for both `QP` and `Scope` by item. Although the interaction term was initially included in the random-effects structure, it was removed due to convergence issues. The final converged model is shown in (41).

(41) **Model (Acceptance Rate):**

```
glmer(Acceptance ~ QP * Scope + (1 + QP | Participants)
+ (1 + QP + Scope | Item), family = binomial
```

Although the mean acceptance rates varied numerically across conditions, these differences did not reach statistical significance in the mixed-effects model. Specifically, the model revealed no significant effect of `QP` ($\beta = 0.09$, SE = 0.36, $z = 0.25$, $p = .800$), no effect of `Scope` ($\beta = 0.01$, SE = 0.38, $z = 0.02$, $p = .982$), and no interaction between the two ($\beta = 0.17$, SE = 0.39, $z = 0.43$, $p = .671$).

These results support the conclusion that scope ambiguity was not systematically resolved in favor of one interpretation across syntactic configurations. Both $neg > \forall$ and $\forall > neg$ readings appear to be equally accessible, regardless of whether the quantifier occurred in subject or object position.

Following the acceptance rate analysis, I examined participants' reaction times (RTs) to assess whether processing difficulty differed across conditions. First, I compared RTs for target items against those for filler sentences to determine whether scope-related processing incurred greater cognitive effort.

`QP`	`Scope`	Mean Reaction Time (ms)
Subject QP	$neg > \forall$	5862
	$\forall > neg$	7453
Object QP	$neg > \forall$	6319
	$\forall > neg$	5778

Table 6: Mean reaction times across conditions

The Shapiro–Wilk test revealed that RT differences were non-normally distributed ($W = 0.65$, $p < .001$). A Wilcoxon signed-rank test confirmed that RTs for test conditions were significantly longer than for filler sentences ($V = 424$, $p < .001$), as visualized in Figure 6. This suggests that

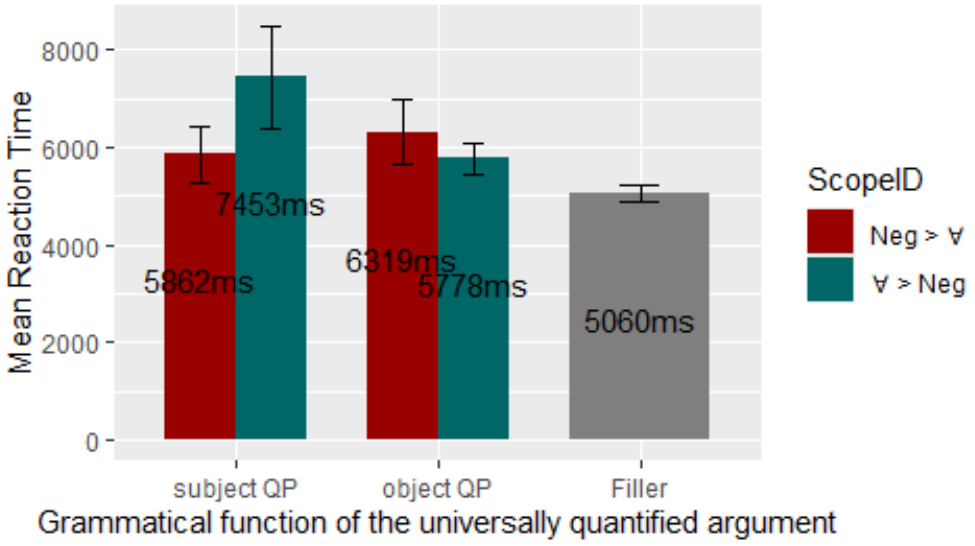


Figure 6: Mean reaction times across conditions

interpreting scope ambiguity imposed greater processing demands than structurally simpler filler items.

To determine whether QP type or scope configuration modulated RTs, I fitted a Linear Mixed-Effects Model (LMM) to log-transformed RTs using the `lmerTest` package in R. The fixed effects mirrored those in the GLMM: `QP`, `Scope`, and their interaction. Random intercepts and slopes were included for both participants and items. The final model is shown in (42).

(42) **Model log-transformed (Reaction Time):**

```
glmer(LogRT ~ QP * Scope + (1 + QP | Participants)
+ (1 + QP * Scope | Item)
```

None of the fixed effects were statistically significant. There was virtually no difference in RTs between subject and object QP conditions ($\beta = -0.001$, SE = 0.055, $t = -0.014$, $p = .989$), and the contrast between *neg* $>$ ∀ and ∀ $>$ *neg* readings was likewise non-significant ($\beta = 0.020$, SE = 0.084, $t = 0.232$, $p = .817$). The interaction also failed to reach significance ($\beta = 0.063$, SE = 0.070, $t = 0.902$, $p = .379$).

Taken together, these results indicate that while scope-related processing requires more effort than interpreting filler items, the processing cost was not reliably modulated by syntactic position or scope configuration. This supports the conclusion that *neg* $>$ ∀ and ∀ $>$ *neg* readings were comparably accessible and that participants did not experience differential difficulty based on structural properties of the sentence.

4 Discussion

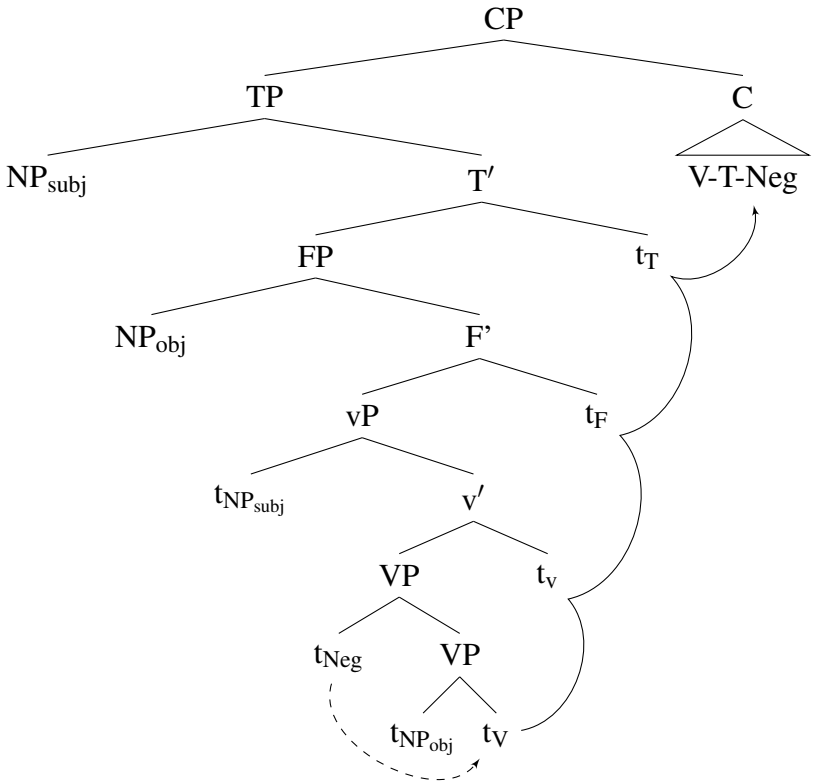
In the previous section, the results showed that both *neg* $>$ ∀ and ∀ $>$ *neg* readings were robustly attested for negative sentences with quantified phrases, regardless of the position of the universal

quantifier *b"uh* 'every.' Both acceptance rates and reaction times indicated that *neg* $>$ $\forall$ and $\forall >$ *neg* readings were equally accessible, rendering the test sentences truly ambiguous.

At first glance, these findings appear paradoxical from the standpoint of V-raising. In Section 2.3, I proposed that Mongolian exhibits rigid scope: scope interpretation should reflect the c-command relationship between the quantifier and negation. Yet we observed the availability of both readings. Given that the relative scope is determined by the position of negation—contingent on V-raising—this ambiguity seems to imply that V-raising is optional. Is that a tenable conclusion?

I argue that the availability of both readings does not result from different syntactic derivations. Instead, I propose that both readings are derived from a single structure in which the verbal complex undergoes V-to-C movement. As a result, the complex verbal head containing negation c-commands both the subject and the object, as shown in (43). This structural configuration uniformly gives rise to the *neg* $>$ $\forall$ reading, regardless of the grammatical function of the quantified phrase. The $\forall >$ *neg* reading, in turn, becomes available by virtue of logical entailment from the *neg* $>$ $\forall$ reading.

(43)



There is a well-established entailment relation between the two readings: whenever the $\forall >$ *neg* reading is true, the *neg* $>$ $\forall$ reading must also be true. For example, consider the sentence 'Every fox did not chase a marmot.' Under the $\forall >$ *neg* reading, it is true when none of the foxes chased a marmot. But this scenario also makes the sentence true under the *neg* $>$ $\forall$ reading, since it rules out the case in which all foxes performed the action—see Table 7. Therefore, when participants judge the sentence to be true in such a case, we cannot determine which reading drove their response. In contrast, when the sentence is judged true in a context where some but not all foxes chased a marmot, this necessarily reflects the *neg* $>$ $\forall$ reading, since the $\forall >$ *neg* reading would be false in

that situation.

Number of the foxes chasing a marmot	Truth value	
	neg $>$ $\forall$	$\forall$ $>$ *neg*
0	T	T
1	T	F
2	T	F
3 (all)	F	F

Table 7: Truth value of 'Every fox did not chase a marmot.'

Given this entailment pattern, Scontras et al. (2017) argue that only the reading that is entailed by the other can provide meaningful insights into scope ambiguity and syntactic structure. Applied to the present study, this means that only the availability of the *neg* $>$ $\forall$ reading allows us to infer the position of negation. Even though the $\forall$ $>$ *neg* reading is not derived from surface structure in this analysis, it remains accessible via logical inference.

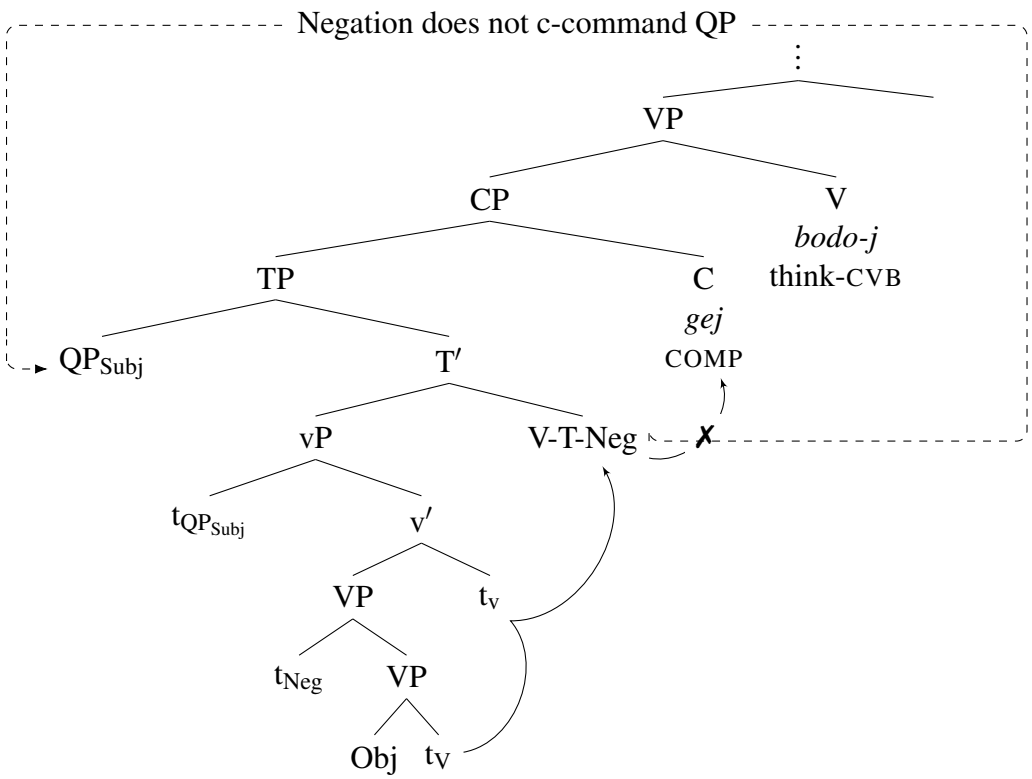
This proposal aligns with the findings in Han et al.'s (2007) analysis of Korean. In their study, the object QP condition yielded a striking split: participants either always accepted the *neg* $>$ $\forall$ reading or never did. This bimodal pattern supported the existence of two speaker groups—those with and without V-raising. For V-raising speakers, negation c-commands the object, yielding the *neg* $>$ $\forall$ reading and, via entailment, the $\forall$ $>$ *neg* reading as well. For non-V-raising speakers, the verb remains below the object, permitting only the $\forall$ $>$ *neg* reading. In this sense, their findings provide empirical precedent for the entailment-based account proposed here.

Turning to the Rich Agreement Hypothesis (RAH), the current results challenge its core prediction. As discussed in Section 2.2, the RAH links V-raising to the presence of morphologically rich agreement. Since Khalkha Mongolian lacks overt agreement in person or number, the RAH predicts that V-raising should not occur. However, the experimental results support an analysis in which the verb raises to C, contradicting the prediction of the RAH. This suggests that V-raising in Mongolian must be driven by other factors.

The proposal that the verb raises to C further predicts that no other overt element should occupy C. If C is filled by another constituent, such as an overt complementizer, V-to-C movement should be blocked, thereby preventing the *neg* $>$ $\forall$ reading. Mongolian provides an ideal test case for this: the complementizer *gej* obligatorily introduces indirect speech, as shown in (44). In such cases, the verb cannot move to C. The prediction, then, is that the *neg* $>$ $\forall$ reading will be unavailable in subject QP conditions under indirect speech—see (45). Assuming V-to-C movement proceeds cyclically, it will stop at T, leaving the verb below the subject QP and yielding only the $\forall$ $>$ *neg* reading.

(44) *Bi* [$_{CP}$ *Altantsetseg-iig zarimdaa Geriin daalgawr-aa khii-deg-**güi** **gej***] *bodoj*
I A.-ACC sometimes home.GEN work-REFL.POSS do-HAB-**not** **COMP** think.CVB
baina.
COP.NPST
'I'm thinking that A. sometimes does not do herself's homework.'
(adopted from Gong 2022: 91)

(45)



To summarize, I argue that scope ambiguity in Mongolian arises not from optionality in verb movement but from a fixed V-to-C structure coupled with logical entailment. Nevertheless, I acknowledge an alternative analysis: one in which the $\forall > neg$ reading arises from true structural ambiguity, i.e., optionality in the landing site of V-raising. Under this account, different configurations yield different readings—e.g., V-to-T yielding $\forall > neg$ scope in subject QP conditions. Optional verb stranding or movement to intermediate heads could also account for both readings, and such optionality is known to exist, for instance, with past participles in some Romance languages (Cinque 1999). Although the current data favor an entailment-based account, nothing here conclusively rules out optional head movement.

Future research can help adjudicate between these two analyses. Negative sentences with numeral quantifiers may provide a crucial test case[5]. Unlike universal quantifiers, numeral quantifiers and negation do not stand in an entailment relation. For example, in a sentence like (46) with a subject quantified as 'two foxes,' the two scope readings yield distinct, non-overlapping truth conditions. Under the $neg > 2$ reading, the sentence is true unless exactly two foxes chased a marmot. Under the $2 > neg$ reading, the sentence is true only if two foxes did not chase a marmot.

(46) Two foxes didn't chase a marmot.

a. Reading 1: It is not the case that two foxes chased a marmot. ($neg > 2$)

b. Reading 2: There are two foxes that didn't chase a marmot. ($2 > neg$)

Suppose we construct two scenarios with a Mongolian counterpart of (46). In the first, four foxes are present—two chased a marmot, two did not. This makes the sentence true under the $2 >$

[5]I thank Chris Kennedy and Erik Zyman for encouraging this line of inquiry.

neg reading but false under the *neg* $>$ 2 reading. In the second scenario, no foxes chased a marmot, so the sentence is true under the *neg* $>$ 2 reading but false under the 2 $>$ *neg* reading. These possibilities are summarized in Table 8. Crucially, since no entailment holds between the two readings, the availability of both cannot be explained via logical inference. If the V-to-C account is correct, only the *neg* $>$ 2 reading should be available, and the sentence should be judged true only in the second scenario. If both readings are available, this would support a model in which the verb optionally raises to different positions depending on discourse context.

The number of the foxes that		2 $>$ *neg*	*neg* $>$ 2
chased a marmot	didn't chase a marmot		
2	2	T	F
1	1	F	T

Table 8: Truth value of (46)

5 Conclusion

In this paper, I have argued that verb raising occurs in Khalkha Mongolian. Using scope interpretation between negation and universally quantified arguments as a diagnostic, I demonstrated that negative sentences containing universal quantifiers are systematically ambiguous between the *neg* $>\forall$ and $\forall >$ *neg* scope readings. Assuming that scope interpretation reflects the asymmetric c-command relationship between the position of negation and that of the quantifier, the availability of the *neg* $>\forall$ reading in both subject and object QP configurations indicates that the negation occupies a structurally high position—one that c-commands both subject and object QPs.

I further proposed that the apparent availability of the $\forall >$ *neg* reading can be attributed to a logical entailment from the *neg* $>\forall$ reading. That is, the $\forall >$ *neg* interpretation does not require a distinct syntactic derivation but arises from the truth-conditional inclusion relation between the two readings. This analysis supports a unified structure in which the verb undergoes V-to-C movement in Khalkha Mongolian, with scope ambiguity resulting from interpretive rather than syntactic variation.

Abbreviations

ABL	ablative
ACC	accusative
COMP	complementizer
COND	conditional
COP	copula
CVB	converb
DAT	dative

DECL declarative
DL Dative/locative
FOC focus
GEN genitive
HAB Habitual
IMPF Imperfective
NEG negative
NOM nominative
NPST Non-past
PL plural
POSS possessive
PST past
PTCP participle
PTL Particle
REFL reflexive
TOP topic

References

Aoun, J. and Y.-h. A. Li (1989). Scope and constituency. *Linguistic inquiry 20*(2), 141–172.

Aoun, J. and Y.-h. A. Li (2003). *Essays on the Representational and Derivational Nature of Grammar: The Diversity of Wh-Constructions*. MIT press.

Bao, L., M. Hasebe, and H. Maki (2015). The negation markers in modern Inner Mongolian. In *Essays on Mongolian Syntax*, pp. 81–99. Tokyo: Kaitakusha Col, Ltd.

Binnick, R. I. (2011). *The Past Tenses of the Mongolian Verb: Meaning and Use*, Volume 1. Brill.

Bobaljik, J. D. and H. Thráinsson (1998, April). Two Heads Aren't Always Better Than One. *Syntax 1*(1), 37–71.

Cho, Y.-M. Y. and K.-S. Hong (1988). Evidence for the VP Constituent from Child Korean.

Choi, J. and H. Harley (2019). Locality Domains and Morphological Rules: Phases, Heads, Node-Sprouting and Suppletion in Korean Honorification. *Natural Language & Linguistic Theory 37*(4), 1319–1365.

Choi, Y.-S. (1999). Negation, Its Scope and NPI Licensing in Korean. *Proceedings - Eastern States Conference on Linguistics (ESCOL) 15*, 25–36.

Chung, D. and H.-K. Park (1997). NPIs outside of Negation Scope. In *Japanese/Korean Linguistics*, Volume 6, pp. 415–435.

Chung, I. (2009). Suppletive verbal morphology in Korean and the mechanism of vocabulary insertion1. *Journal of Linguistics 45*(3), 533–567.

Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press.

Crain, S. and R. Thornton (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Language, Speech, and Communication. Cambridge, Mass: MIT Press.

Dagvasumberel, E. (2015). Sentential Negative Markers as Pro-Forms of Negative Sentences in Modern Mongolian. *Advances in Language and Literary Studies 6*(1), 87–93.

Folli, R. and H. Harley (2005). Flavors of v: Consuming Results in Italian & English. In P. Kempchinsky and R. Slabakova (Eds.), *Aspectual Inquiries*, Volume 62, pp. 95–120. Dordrecht: Springer Netherlands.

Fukui, N. and H. Sakai (2003). The Visibility Guideline for Functional Categories: Verb Raising in Japanese and Related Issues. *Lingua 113*(4-6), 321–375.

Giorgi, A. and F. Pianesi (1997). *Tense and Aspect: From Semantics to Morphosyntax*. Oxford University Press.

Gong, Z. M. (2022). *Issues in the Syntax of Movement: Cross-clausal Dependencies, Reconstruction, and Movement Typology*. Cornell University.

Guntsetseg, D. (2012). *Differential Case Marking in Mongolian*. Ph. D. thesis, University of Stuttgart.

Guntsetseg, D. (2016). *Differential Case Marking in Mongolian* (1 ed.). Harrassowitz Verlag.

Han, C.-h. and C. Lee (2007). On negative imperatives in Korean. *Linguistic Inquiry 38*(2), 373–395.

Han, C.-h., J. Lidz, and J. Musolino (2007). V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry 38*(1), 1–47.

Heim, I. (1983). On the projection problem for presuppositions. *Formal semantics–the essential readings*, 249–260.

Hoji, H. (1998). Null object and sloppy identity in Japanese. *Linguistic inquiry 29*(1), 127–152.

Huang, C. T. J. (1982). Move WH in a Language without WH Movement. *The Linguistic Review 1*(4).

Huang, S.-F. (1981). On the scope phenomena of Chinese quantifiers. *Journal of Chinese Linguistics 9*(2), 226–243.

Janhunen, J. (2006). Khamnigan Mongol. In J. Janhunen (Ed.), *The Mongolic Languages*, Routledge Language Family Series. London & NY: New York: Routledge.

Kim, S. (1999). Sloppy/Strict Identity, Empty Objects, and NP Ellipsis. *Journal of East Asian Linguistics 8*(4), 255–284.

Klima, E. (1964). Negation in English. In J. A. Fodor and J. Katz (Eds.), *The Structure of Language*, pp. 246–323. NJ: Englewood Cliffs: Prentice-Hall.

Koeneman, O. and H. Zeijlstra (2014). The rich agreement hypothesis rehabilitated. *Linguistic Inquiry 45*(4), 571–615.

Koisumi, M. (2000). String vacuous overt verb-raising. *Journal of East Asian Linguistics 9*(3), 227–285.

Kullmann, R. and D. Tserenpil (2008). Mongolian Grammar. Ulaanbaatar: Admon Co.

Kuroda, S.-Y. (1965). *Generative Grammatical Studies in the Japanese Language.* Ph. D. thesis, Massachusetts Institute of Technology.

Lee, T. H.-t. (1986). *Studies on Quantification in Chinese*. Ph. D. thesis, University of California, Los Angeles.

Lee, T. T.-M. (2023, December). Last but not least: A comparative perspective on right dislocation in Alasha Mongolian. *Journal of East Asian Linguistics 32*(4), 459–495.

Lidz, J. and J. Musolino (2002). Children's command of quantification. *Cognition 84*(2), 113–154.

Lim, J. J. (2023, September). Finiteness, agreement and subject case in two varieties of Mongolian.

Lubsangdorji, J. and J. Vacek (2004). *Colloquial Mongolian: An Introductory Intensive Course*, Volume 1. Prague: Triton & Charles University.

Miyagawa, S. (2001). The EPP, scrambling, and wh-in-situ. In M. Kenstowicz (Ed.), *Ken Hale: A Life in Language*, Volume 36, pp. 293–338. MA: Cambridge, London: The MIT Press.

OpenAI (2024). ChatGPT. OpenAI.

Otani, K. and J. Whitman (1991). V-raising and VP-ellipsis. *Linguistic Inquiry*, 345–358.

Peters, A. (2020). Scrambling for case: Accusative in Mongolian. In *Proceedings of the 2020 Annual Conference of the Canadian Linguistic Association*, pp. 1–15.

Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic inquiry 20*(3), 365–424.

Pullum, G. K. and A. M. Zwicky (1983). Cliticization vs. inflection: English n't. *Language 59*(3), 502–513.

Scontras, G., M. Polinsky, C.-Y. E. Tsai, and K. Mai (2017). Cross-linguistic scope ambiguity: When two systems meet. *Glossa: A journal of general linguistics 2*(1), 1–28.

Slater, K. W. (2005). *A Grammar of Mangghuer: A Mongolic Language of China's Qinghai-Gansu Sprachbund*. Routledge.

Velupillai, V. (2012). An introduction to linguistic typology.

Yoon, J. H. (1994). Korean Verbal Inflection and Checking Theory. *MIT Working Papers in Linguistics 22*(July), 251–270.

Zhou, P. and L. Gao (2009). Scope processing in Chinese. *Journal of Psycholinguistic Research 38*, 11–24.

Appendix: The list of the items used for sentence generation