Variation in negation and quantifier scope judgments in Korean∗

Chung-hye Han, Jeffrey Lidz and Dennis Ryan Storosheko
Simon Fraser University, Maryland University, and Yale University

1 Introduction
Korean has two forms of negation: long negation is post-verbal and requires ha to support tense and other verbal inflections, as in (1), while short negation is preverbal, with no ha-support, as in (2). Both forms are clitics and must attach to the main verb of the sentence.

(1) Dora-ka khwukhi-ul mek-ci ani ha-yess-ta.
   Dora-NOM cookie-ACC eat-CONN NEG do-PST-DECL
   ‘Dora did not eat a cookie.’

(2) Dora-ka khwukhi-lul an mek-ess-ta.
   Dora-NOM cookie-ACC NEG eat-PST-DECL
   ‘Dora did not eat a cookie.’

Linguists have used judgments on scope interaction between object quantified phrase (QP) and negation to determine the syntactic height of negation and the verb in Korean. But the scope judgments reported in the literature often conflict with one another (Cho 1975; Song 1982; Suh 1989; Hagstrom 2000; Hagstrom 2002; Baek 1998; Kim 2002; Kim 2000), raising doubts as to the usefulness of such data to reach any meaningful conclusions about the clause structure. Han, Lidz and Musolino (2007), however, showed experimentally that the variability in scope judgment exists between speakers and not within a speaker. They thus concluded that there are two groups of speakers of Korean, and attribute this split to the co-existence of two grammars differing on the height of the verb in the population (*the two-grammar hypothesis*): the neg>Q reading is available to speakers with a verb-raising grammar but unavailable to those with an INFL-lowering grammar.

The two-grammar hypothesis makes two additional predictions. First, an individual speaker should exhibit the same scope judgments for both types of negation. This prediction follows because for both constructions the height of the verb determines the scope of negation in the same way. Second, if an individual speaker maintains only one grammar, then he should exhibit the same scope judgment over time. In this paper, we present an experiment that tests these predictions. It will be shown that both predictions are borne out.

This paper is organized as follows. In section 2, we discuss why scope of negation and argument QPs is a good test for verb-raising in Korean. In section 3, we discuss Han, Lidz and Musolino’s (HLM) experimental findings and the predictions they make. We then present our experiment and findings, and their implications for the two-grammar hypothesis in section 4.

∗We thank the audience at the CLS 47 for insightful comments and questions. This work was partially supported by SSHRC 410-2007-2169.
2 Scope of negation and argument QPs as a test for verb-raising

Korean is verb-final. So, any diagnostic test based on word order cannot be used to tell us the syntactic height of the verb: whether the verb raises or not, it will be at the end of the sentence. Instead, as observed in HLM, scope of negation and argument QPs can be informative because of three background facts about Korean: frozen scope, object raising and clitic status of negation.

First, in Korean, argument QPs exhibit frozen scope: in a sentence as in (3) with canonical SOV word order, with subject and object QPs, the only reading available is the one in which the subject takes scope over the object (Joo 1989; Sohn 1995; Hagstrom 2000).

(3) **Nwukwunka-ka manhun salam-ul piphanhay-ss-ta.**
    someone-NOM many person-ACC criticize-PST-DECL
    ‘Someone criticized many people.’
    (some > many, * many > some)

Second, the object phrase in a transitive sentence must occur to the left of (i.e., higher than) vP/VP adverbs, such as *cal ‘well’, as in (4). This provides support for the view that objects raise from a VP-internal position to a function projection higher in the clause structure (Hagstrom 2000; Hagstrom 2002).

(4) a. Toli-ka **maykcwu-lul cal** masi-n-ta.  (S O Adv V)
    Toli-NOM beer-ACC well drink-PRES-DECL
    ‘Toli drinks beer well.’

    b. * Toli-ka **cal maykcwu-lul** masi-n-ta.  (*S Adv O V)
    Toli-NOM well beer-ACC drink-PRES-DECL
    ‘Toli drinks beer well.’

Third, both long and short negation have the morphosyntactic status of clitics, and are treated as a unit with the verb in the overt syntax. Long negation must occur immediately before *ha as in (5), and short negation must occur immediately before the lexical verb as in (6).

(5) a. Toli-ka **maykcwu-lul cal ani ha-n-ta.**
    Toli-NOM beer-ACC well drink-CONN NEG do-PRES-DECL
    ‘Toli doesn’t drink beer well.’

    b. * Toli-ka maykcwu-lul masi-ci **ani cal ha-n-ta.**
    Toli-NOM beer-ACC drink-CONN NEG well do-PRES-DECL
    ‘Toli doesn’t drink beer well.’

(6) a. Toli-ka **maykcwu-lul an masi-n-ta.**
    Toli-NOM beer-ACC NEG drink-PRES-DECL
    ‘Toli doesn’t drink beer.’

    b. * Toli-ka **an maykcwu-lul masi-n-ta.**
    Toli-NOM NEG beer-ACC drink-PRES-DECL
    ‘Toli doesn’t drink beer.’
Given the scope freezing effect, the scope of argument QPs will be determined in their surface position, without recourse to quantifier raising or reconstruction. This then means that the relative scope of negation and an argument QP will be determined by the position of negation in the clause structure. Finally, given that objects must raise out of the vP/VP and that negation is a unit with the verb, the relative scope of negation and an object QP will tell us whether the verb has raised. If the verb raises, then negation will occur in a position higher than an object QP and will therefore take scope over this QP. But if the verb remains low, then negation will also remain low and the object QP will take scope over negation.

These predictions can be fleshed out in more detail with the clause structure postulated in HLM.\(^1\) In (7a), long negation heads its own projection, and ha-support is assumed to take place in F. In (7b), short negation is adjoined to the left of VP. In both structures, the subject is in [Spec,IP], and the object raises to a functional projection (FP) external to VP. As the subject is high in the structure, a subject QP will scope over negation, independent of negation type. If ha raises in (7a) and verb raises in (7b), then long negation cliticized to ha and short negation cliticized to the verb will occur in IP and so they will take scope over object QPs. If there is no ha-raising or verb-raising, then both long and short negation will remain below [Spec,FP] and so object QPs will scope over it.

The predictions are clear, but there is much disagreement in the literature as to what the facts are. In the next section, we present HLM’s study that found that there is indeed variation in scope judgments among the speakers of Korean and that the judgments vary not within a speaker but between speakers.

3 Han, Lidz and Musolino’s (2007) findings

HLM used the Truth Value Judgment Task (Crain & Thornton 1998) in an experiment to test scope judgments of sentences containing a subject universal QP and

\(^1\)The structures in (7a) and (7b) abstract away from splitting the verbal projection into vP and VP and raising the subject from a lower position to [Spec,IP].
long negation as in (8a), a subject universal QP and short negation as in (8b), an object universal QP and long negation as in (9a) and an object universal QP and short negation as in (9b).

(8) Subject QPs
      every horse-NOM fence-ACC jump.over-CI NEG do-PST-DECL
      ‘Every horse didn’t jump over the fence.’ (long negation)
      every horse-NOM fence-ACC NEG jump.over-PST-DECL
      ‘Every horse didn’t jump over the fence.’ (short negation)

(9) Object QPs
      cookie monster-NOM every cookie-ACC eat-CI NEG do-PST-DECL
      ‘Cookie monster didn’t eat every cookie.’ (long negation)
      cookie monster-NOM every cookie-ACC NEG eat-PST-DECL
      ‘Cookie monster didn’t eat every cookie.’ (short negation)

The experiment had a between-subjects design, and tested 3 factors with 2 levels each: scope (neg>every vs. every>neg) × negation type (short vs. long) × grammatical function of the QP (subject vs. object). It was thus divided into 8 different conditions to which 8 different sets of 20 participants were randomly assigned. In a trial, a participant watched a short video clip of an experimenter enacting a scenario using toys followed by a puppet making a statement (test sentence) about the scenario. The participant’s task was to indicate whether the puppet’s statement is true or not. For example, in the scenarios that tested the neg>every reading with test sentences in (8) and (9), two out of three horses (i.e., not all horses) jumped over the fence, and two out of three cookies (i.e., not all cookies) were eaten. In the scenarios that tested the every>neg reading with (8) and (9), none of the horses jumped over the fence and none of the cookies were eaten. Each participant was given 4 test trials and 4 filler trials in a pseudo-random order.

Figures 1 and 2 summarize mean percentage acceptances by condition. HLM found (i) a main effect for scope ($p<.0001$), and (ii) a main effect for grammatical function ($p<.0008$) and an interaction between scope and grammatical function ($p<.0003$). So, regardless of negation type or grammatical function, speakers were more likely to accept the every>neg reading than the neg>every reading, and independently of negation type, speakers were more likely to accept the neg>every reading on an object QP than they were on a subject QP.

With subject QPs, speakers in general accepted the every>neg reading and rejected the neg>every reading. In Figure 1, while the acceptance rate on the every>neg reading is 100% for both short and long negation, the acceptance rate on the neg>every is only 4% for short negation and 19% for long negation. This confirms the prediction laid out at the end of section 2 that subject QPs will scope
over negation, independent of negation type. But speakers seem to be divided when it comes to the neg>every reading with object QPs: although the acceptance rate on the neg>every reading was higher in object conditions than in subject conditions, 37% with short negation and 46% with long negation, over 50% of the participants still did not accept this interpretation in object conditions. Upon further analysis of the data, HLM found that speakers tend to exhibit a bimodal distribution of responses in the neg>every context in object conditions. That is, speakers either almost always accepted or almost always rejected the neg>every reading. This tendency is shown in Figure 3, which divides the participants into groups based on their rate of acceptance of the neg>every reading in object conditions.

HLM argue that this bimodal distribution in responses is a reflection of two grammars in competition in the speech community of Korean: those speakers who have acquired the grammar with verb-raising and ha-raising accept the neg>every reading with object QPs, and those who have acquired the grammar with INFL-lowering reject the neg>every reading. The reason for the co-existence of two grammars, they argue, is that children acquiring Korean, being head-final, are unlikely to receive sufficient input that provides clear evidence for or against the syntactic height of the verb.

This two-grammar hypothesis raises two questions, which are unanswered in HLM. First, do speakers of Korean show the same pattern of behaviour for both long and short negation? This is predicted to be so because for both types of negation, the height of the verb that hosts negation determines the scope of negation in the same way. HLM’s study showed roughly the same size split in the population.

Figure 1: Mean Percentage Acceptance: Subject

Figure 2: Mean Percentage Acceptance: Object
for both long and short negation, but it doesn’t tell us whether any given individual would show the same pattern for both types of negation because their experiment had no participants who were tested on both types of negation. Second, is the split population observed in HLM’s study due to each speaker stably controlling one grammar or to each speaker oscillating between both grammars? HLM hypothesized that a speaker controls a single grammar. However there is an alternative interpretation of their result in which each individual maintains both a verb-raising and an INFL-lowering grammar. This is because in the HLM’s experimental task, it may be that the grammar that was chosen on the first item, whichever that is, may be exerting an influence over subsequent items, priming the participants’ answers. This kind of priming effect would give the appearance of two populations of speakers when in fact there was a single population in which each speaker controlled two competing grammars. We address these questions in our experiment described in the next section.

4 Our experiment
4.1 Design
Using the Truth Value Judgment Task similar to HLM in a within-subjects experiment, adult speakers of Korean were tested on the scope of negation and object QP on two separate occasions, 1 month apart. We only tested sentences with object QP and not subject QP, as these are potentially informative about the height of the verb. The experiment tested 3 factors with 2 levels each: negation type (short vs. long) × scope (every>neg vs. neg>every) × test session (March vs. April). The experiment was thus divided into 8 conditions. As this is a within-subjects experiment, each participant was tested on all 8 conditions. Test sentences were similar to those used in HLM, containing an object QP and long or short negation, but they contained an adverbial phrase right after the object, as in (10) and (11). The reason for adding the adverbial phrase to test sentences was to further ensure the object raising to a vP/VP external position.
(10) Object QP and short negation

Acessi-ka motun catongcha-lul cip ap-eyse an ssis-ess-ta
man-NOM every car-ACC house front-in NEG wash-PST-DECL
‘The man did not wash every car in front of his house.’

(11) Object QP and long negation

Totwuk-i motun posek-ul kakey-eyse hwumchici ani ha-yess-ta.
burglar-NOM every jewel-ACC store-at steal NEG do-PST-DECL
‘The burglar did not steal every jewel in the store.’

31 participants were tested in the March session, and from these 26 participants were tested again in the April session. The remaining 5 participants from the March session chose not to participate in the April session. The participants were introduced to the task with 4 practice trials, then were given a set of 16 test trials in the first session (4 trials per each scope/negation combination), and a different set of 16 test trials in the second. The same set of 12 fillers were used in both sessions.

4.2 Findings

Figure 4 summarizes mean percentage acceptances by condition. We found a main effect of scope \((p<.001)\), with participants more likely to accept the every\(>\)neg reading than neg\(>\)every reading, but no other main effects or interactions. This suggests that the speakers behaved uniformly across negation types and test sessions: those who rejected the neg\(>\)every reading did so on both test sessions and for both types of negation, and those who accepted the neg\(>\)every reading did so on both test sessions and for both types of negation.

![Figure 4: Mean Percentage Acceptances by Condition](image)

To confirm the uniform behaviour of each individual participant across negation types, we calculated, per each participant, the difference score between the acceptance rate in short negation/neg\(>\)every condition and the acceptance rate in long negation/neg\(>\)every condition for both March and April sessions. A negative difference score indicates that a participant was more likely to accept the neg\(>\)every
interpretation for long negation than short negation and a positive difference score indicates that a participant was more likely to accept the neg\textgreater every interpretation for short negation than long negation. A difference score of 0 means that a participant behaved the same across negation types. Figure 5 plots the count of difference scores. The figure shows that the majority of participants behaved the same across negation types in both March and April.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure5}
\caption{Number of participants in each difference score between Short neg and Long neg in neg\textgreater every conditions: March (top), April (bottom)}
\end{figure}

In the march session, the average difference score between short and long negation was -0.08 (se = 0.03). Participants were slightly more likely to accept the neg\textgreater every interpretation in long-negation over short negation ($t(30)=-3.24$, $p<.003$), as shown in Figure 5 (top). This significant difference derives from the fact that 10/31 participants gave one more “yes” response for long negation than short negation, though only one participant gave one more “yes” response for short negation than long negation. However, because the difference scores are so close to 0 in all cases, we do not take this to show that people are inconsistent in their choice of verb-raising across long and short negation. Further evidence for this conclusion comes from the April data, where the mean difference score between long and short
negation was -0.02 (se = 0.02), as shown in Figure 5 (bottom). This difference score was not significantly different from zero ($t(30)=-0.9$, ns), lending further support to the view that people maintain a single grammar of verb-raising for both long and short negation.

Turning now to the potential effect of test-session, we find no significant effects, consistent with the view that each subject controls only a single grammar. We calculated, per each participant, the difference score between the acceptance rate of March/neg every condition and April/neg every condition for both short negation and long negation. Figure 6, which plots the count of difference scores, shows that the majority of participants behaved the same across test sessions for both short negation and long negation.

**Figure 6:** Number of participants in each difference score between March and April in neg every conditions: Short neg (top), Long neg (bottom)

For short negation (Figure 6 (top)), the mean difference score between March and April was -0.03 (se=0.06). This is not significantly different from 0 ($t(25)=-0.64$, ns). Thus, we have no evidence that participants changed judgments across test sessions. For long negation (Figure 6 (bottom)), the mean difference score between March and April was 0.01 (se = 0.05). This is not significantly different
from 0 ($t(25)=0.35, \text{ns}$). Again, we have no evidence that participants changed judgments across test sessions.

4.3 Discussion

We can now answer the two questions raised at the end of section 3 about the two-grammar hypothesis of HLM. Do speakers of Korean show the same pattern of behaviour for both long and short negation and is the split population observed in HLM’s study due to each speaker stably controlling one grammar? According to our findings, the answer to both questions is ‘yes’: a given individual shows the same pattern for both types of negation and he maintains one grammar and does not oscillate between two grammars over time.

Both results are predicted by the two-grammar hypothesis, and thus provide further support for the proposal advanced by HLM. In the population that has not acquired a verb-raising grammar, the neg$>\text{every}$ reading for an object QP in sentences formed with short negation is not available because the grammar only generates the structure in which the object scopes over negation. In the same population, ha-raising does not take place, and so neg$>\text{every}$ reading for an object QP is also unavailable in sentences with long negation. On the other hand, the population that has acquired a verb-raising grammar generates the neg$>\text{every}$ reading for an object QP in sentences formed with short negation because the grammar generates the structure in which negation scopes over the object. In the same population, in sentence with long negation, ha-raising will take place, again generating the neg$>\text{every}$ reading for an object QP. The grammar with no verb-raising and no ha-raising is represented as Grammar A in (12), and the grammar with verb-raising and ha-raising is represented as Grammar B in (13).

(12) Grammar A

<table>
<thead>
<tr>
<th>IP</th>
<th>NP$_{Subj}$</th>
<th>FP</th>
<th>I</th>
<th>NP$_{Obj}$</th>
<th>VP</th>
<th>F</th>
<th>ShNeg</th>
<th>NP</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Infl lowers to V; Short neg cliticizes to V; Object scopes over short neg.

<table>
<thead>
<tr>
<th>IP</th>
<th>NP$_{Subj}$</th>
<th>FP</th>
<th>I</th>
<th>NP$_{Obj}$</th>
<th>NegP</th>
<th>F</th>
<th>Neg</th>
<th>NP</th>
<th>V</th>
<th>LongNeg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Infl lowers to F; Long neg cliticizes to F; Object scopes over long neg.
One puzzle that is raised by our within-subjects experiment has to do with the high acceptance rates for the neg > every reading, in comparison to those in HLM’s between-subjects experiment. In HLM, the mean percentage acceptances in neg > every/object conditions for short negation and long negation were 37% and 46% respectively, whereas in the current within-subjects experiment, they range from 73% to 81%. Though we do not have a good explanation as to why we obtained such high acceptance rates in our within-subjects experiment, we would like to re-emphasize the fact that we found significant difference between neg > every and every > neg, just as in HLM. Another fact worth noting is that Lee et al. (To appear) reports 54.6% as an acceptance rate of neg > every reading for short negation in a within-subjects TVJT experiment (long negation was not tested). This acceptance rate is consistent with the overall proposal of the two-grammar hypothesis.

5 Conclusion
The evidence presented above supports the view that the primary linguistic data that Korean learning children are exposed to is not sufficient to guarantee uniform convergence to a single grammar. The data that would allow a child to choose either verb-raising or INFL-lowering is simply not available. Consequently, it seems that learners choose a single grammar at random and discard the other option. Because there is no basis for this choice, we find some learners choosing one grammar and others choosing another. The consistency within an individual that we find across testing sessions and across negation types lends further support to this view. If speakers maintained both grammars simultaneously, then we would expect them to fluctuate in which grammar they chose to use for any given sentence. Because we
did not see such fluctuation, we can be confident that learners maintain only one grammar.

References


