Bilingualism: Language and Cognition

cambridge.org/bil

Children's interpretation of negation and quantifier scope in L3 English

Kyuhee Jo[1], Kitaek Kim[2] and Hyunwoo Kim[3]

[1]Gyeongin National University of Education; [2]Seoul National University and [3]Yonsei University

Research Article

Cite this article: Jo K, Kim K, Kim H (2021). Children's interpretation of negation and quantifier scope in L3 English. *Bilingualism: Language and Cognition* **24**, 427–438. https://doi.org/10.1017/S1366728920000395

Received: 30 September 2019
Revised: 7 June 2020
Accepted: 8 June 2020
First published online: 21 July 2020

Keywords:
child L3 acquisition; negation quantifier scope; crosslinguistic influence; community language proficiency

Address for correspondence:
Kitaek Kim, Email: kitaek@snu.ac.kr

Abstract

Languages differ in the preferences for the interpretation of the scope relation between negation and a quantifier. This study investigates the understudied issue of how interpretive preferences associated with a quantifier scope in learners' L1 and L2 affect their scope interpretations in L3 acquisition. Based on the current models of L3 acquisition, we tested which language, L1 or L2, exerts a stronger effect on the L3 acquisition of quantifier scope. To this end, the study involved two groups of multilingual children (11–13 years old) with different L1s (Chinese or Russian) but with the same L2 (Korean) and L3 (English). The participants completed truth-value judgment tasks designed to investigate their interpretation patterns for English sentences with negation and a quantifier (e.g., *Tom did not cut all the trees*). The results showed that both groups preferred the L3 interpretation similar to that preferred in their L2, but not in their L1, suggesting a potential L2 influence on L3 acquisition. The study evaluates L3 acquisition theories in light of these results.

Introduction

Sentences that give rise to structural ambiguities present significant difficulties for native and nonnative speakers. One illustrative case showing such ambiguity is the interpretation of quantifier scope. In the sentence *Tom did not cut all the trees*, for example, which contains two scope-bearing operators, the negation *not* and the quantified object *all the trees*, two different interpretations are possible, depending on the scope relation between negation and quantifier – that is, which operator initially takes scope over the other (Han, Lidz & Musolino, 2007; O'Grady, Lee & Kwak, 2009a; Zhou & Crain, 2009). When *not* takes scope over *all*, as in the surface syntax, the sentence yields a partitioned-set interpretation, denoting that *Tom cut only some of the trees*; conversely, when *all* takes scope over *not*, this inverse scope (i.e., against the surface syntax) leads to a full-set interpretation,[1] meaning *Tom cut none of the trees*.

Scope ambiguities have been observed across many languages. Notably, speakers of different languages have different preferences in their scope interpretation. For example, native English-speaking adults have a stronger preference for the partitioned-set interpretation of *Tom did not cut all the trees* than for the full-set interpretation. In contrast, native Korean-speaking adults have a stronger propensity for the full-set interpretation for the translation equivalent. This preference difference across languages has attracted huge attention from scholars of L2 acquisition who explore cross-linguistic influence and language transfer (see Chung, 2013; Kwak, 2014 on universal quantifier and negation scope; see Ionin, Luchkina & Stoops, 2014; Marsden, 2009 on double-quantifier scope).[2]

Despite the prolific research on the effect of transfer in L2 scope interpretation, to date no study has examined transfer in scopal interpretations in the context of L3 acquisition. A question arises as to which language, L1 or L2, will more likely affect the interpretation of an ambiguous L3 sentence with two scope-bearing operators. Theories on L3 acquisition have focused on transfer, attempting to identify the primary source language(s) (e.g., L1 and/or L2) affecting L3 acquisition. However, most previous studies to date on L3 acquisition have dealt with a single linguistic domain with a particular focus on morphosyntactic transfer (González Alonso, Rothman, Berndt, Castro & Westergaard, 2017), raising a question as to how transfer operates in the L3 acquisition of target structures involving an interface between multiple domains (Slabakova & García-Mayo, 2015).

To address this issue, the current study focuses on L3 learners' interpretive preference for the scope of negation and a quantifier, which occurs at the interface among syntax, semantics,

[1]The terms partitioned-set interpretation and full-set interpretation are from Lee (2009).

[2]Although cross-linguistic influence and transfer are often used interchangeably in the literature of L2 acquisition, the two concepts need to be distinguished. Transfer refers to "reduplication from previously acquired linguistic representations" (Rothman, González Alonso & Puig-Mayenco, 2019, p. 24) whereas cross-linguistic influence encompasses cross-language interference during processing. As our investigation of L3 scope interpretation concerns linguistic representations rather than a processing domain, we used the term transfer throughout the paper.

CAMBRIDGE UNIVERSITY PRESS

and pragmatics (Zhou & Crain, 2009). To this end, we examined L3-English children's interpretation preferences for English sentences with negation and a quantified object (e.g., *Tom did not cut all the trees*). The children were learning English in Korea; their L2 was Korean, and their L1 was either Chinese or Russian. As will be elaborated in the following section, the four languages investigated in this study, English, Korean, Chinese, and Russian, diverge in scope interpretations. Although Korean, the learners' L2, is more likely to induce full-set interpretation, the other three languages allow for a stronger preference for the partitioned-set interpretation. In this regard, examining the preference of scope interpretations in L3 English by speakers with the same L2 but with different L1 backgrounds provides an ideal scenario to test the relative weighting of learners' L1 and L2 in their L3 English acquisition.

Literature Review

Linguistic background and L2 acquisition studies

Structural ambiguities in sentences involving negation and quantifiers are resolved in different manners across English, Chinese, Russian, and Korean. See (1a-d) for examples.

(1) a. Tom did not cut all the trees. (English)
b. Tāngmǔ méiyǒu kǎn diào suǒyǒu de shùmù. (Chinese)
Tom not cut all of tree
'Tom did not cut all the trees.'

c. Tom ne srubil vse derev'ya. (Russian)
Tom not cut-3SG all trees-PLU
'Tom did not cut all the trees.'

d. Thom-i motun namwu-lul an call-ass-ta. (Korean)
Tom-NOM all tree-ACC not cut-PST-DECL[3]
'Tom did not cut all the trees.'

The co-occurrence of negation and a quantifier (i.e., scope operators) in a sentence can lead to an interpretive ambiguity. In the surface syntax of English, Chinese, and Russian (1a-c), the negation words *not*, *méiyǒu,* and *ne* c-command the quantified objects *all the trees*, *suǒyǒu de shùmù*, and *vse derev'ya*, respectively; in contrast, in the surface syntax of Korean (1d), the quantified object *motun namwu* 'all the trees' c-commands the negation word *an* 'not' (Han et al., 2007, p. 21).

The distinct c-command relations across the languages give rise to the partitioned-set interpretation in English, Chinese, and Russian and the full-set interpretation in Korean. However, if a movement occurs at the level of semantic representation (so-called quantifier raising; May, 1977, 1985), the c-command relation becomes reversed, resulting in the full-set interpretation in English, Chinese, and Russian because the quantified object c-commands the negation word (Lidz & Musolino, 2002). As for Korean, the negation word cliticizes to V and then raises to I, which ultimately c-commands the quantified object (Han et al., 2007).

Despite such potential ambiguity, each language has a rigid interpretation preference, which follows the surface syntax. In Chinese, Russian, and English, there is an interpretation bias for the partitioned-set interpretation (1a-c) (Ionin et al., 2014; Musolino & Lidz, 2003, 2006; O'Grady, 2013; Wu & Ionin, 2019, Zhou & Crain, 2009).[4] In contrast, Korean speakers show a stronger bias for the full-set interpretation (1d) (Han et al., 2007; O'Grady, 2013; O'Grady, Kwak, Lee & Lee, 2011; O'Grady et al., 2009a). Anderson (2004, p. 31) explained this interpretation preference in terms of "processing scope economy," stating that "the human sentence processing mechanism prefers to compute a scope configuration with the simplest syntactic representation (or derivation)." She further mentioned that "computing a more complex configuration is possible but incurs a processing cost." According to this processing scope economy principle, the surface scope interpretation is preferred over the inverse scope interpretation because of the simplicity of the syntactic representation, which does not involve further movement operations and thus requires less processing cost. In contrast, the operation of inverse scope reading involves a movement in the semantic representation (i.e., an extra step) and is therefore taken to be more complex than the surface scope interpretation. Similar to this account, O'Grady (2013) proposed that a language processor prefers processing that does not increase the burden on working memory and thus prefers to assign interpretations that require no revision.[5]

Previous L2 studies of quantifier scope have worked with participants who speak two languages that show contrastive preferences (e.g., English vs. Japanese/Korean). These studies were concerned with L2 learnability issues and tested whether L2 learners can arrive at the target-like interpretation despite the contrastive interpretation in their L1. For example, Marsden (2009) tested L1 English speakers learning Japanese, which disallows inverse scope readings in double-quantifier sentences, and found that intermediate-level learners mistakenly allowed inverse scope readings in Japanese (presumably due to L1 transfer) whereas advanced learners showed target-like interpretations. Chung (2013) tested Korean learners of L2 English with sentences involving negation and quantified objects and found that high proficiency L2 learners showed target-like interpretations (i.e., the partitioned-set interpretation), despite the contrastive preference in the L1 (i.e., the full-set interpretation). Ionin et al. (2014) examined English speakers' scope interpretations in Russian, a language with case marking and free word order. As Russian

[3]3SG = third person singular; PLU = plural; NOM = nominative case; ACC = accusative case; PST = past tense; DECL = declarative marker

[4]However, the interpretive preferences differ between adults and children in some languages. Zhou and Crain (2009) presented that a quantifier c-commanding a negation in the surface syntax results in a full-set interpretation (ia) while a negation c-commanding a quantifier results in a partitioned-set interpretation (ib) in Chinese.

(i) a. Mei-pi ma dou méiyǒu tiaoguo liba.
Every-CL horse all not-have jump-over fence
'Every horse did not jump over the fence.'
b. Bushi mei-pi ma dou tiaoguo-le liba
Not-be every- CL horse all jump-over-ASP fence
'Not every horse jumped over the fence.'

However, Chinese children (3;4–4;3 years old) gave both interpretations for (ia-b) in a TVJT experiment, suggesting that both the surface and inverse scope are available to them. Based on the results, Zhou and Crain claimed that Chinese children start off with a flexible scope relation between quantifier and negation and then shift to rigid interpretation as they get older.

[5]Additional factors may influence scope interpretation, such as the knowledge of entailment relations and the computation of scalar implicature (Chung, 2013; Musolino & Lidz, 2006). For example, when scalar implicature comes into play, a stronger term (which entails the weaker term, but not vice versa) can be less likely preferred; if the speakers intend to deliver a stronger term, they would use a particular sentence explicitly denoting the stronger term, rather than an ambiguous sentence, given that information should be provided as much as required (maxim of quantity; Grice, 1989). The effect of scalar implicature explains a stronger bias toward a partitioned-set interpretation.

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

instantiates the canonical word order of SVO, as in (1c), the strong preference for surface scope readings leads to the partitioned-set interpretation. However, an inverse scope can be accessible with a particular word order, as in scrambled OVS sentences (Ionin et al., 2014), which induce the full-set interpretation. Ionin et al. found that the L2-Russian learners showed the partitioned-set interpretation for the scrambled sentences, a preferred interpretation in their L1 English, which contrasted the full-set interpretation by native Russian speakers in the same sentences. The findings from the previous studies on L2 scope interpretations suggest that L2 learners are likely to carry over their L1 preference for the interpretation of the L2 sentences, particularly when their L2 proficiency is low.

Theories of L3 acquisition

Since Klein's (1995) seminal article about the acquisition of lexical items and syntactic constructions in L3 English, several approaches have attempted to model the transfer of linguistic properties in L3 acquisition. Most models have focused on the initial state of L3 learning and considered which language, L1 or L2, is recruited as the initial grammar for developing the L3 interlanguage. This section reviews some influential ideas and models on L3 acquisition. The first four specify a particular language for the source of transfer (i.e., L1, L2, typologically similar language, and dominant language), while the others propose different languages depending on certain conditions.

Hermas (2010) showed that linguistic properties of the L1 are the only possible source of transfer in L3 processing, maintaining that L1 properties have priority over L2 properties in the initial state of L3 acquisition (Hermas, 2014, 2015, 2018). On the other hand, the L2 Status Factor Model proposed by Bardel and Falk (2007) claims that the L2 acts as the default source of grammatical transfer, for it represents the most recently acquired foreign language. The L2 Status Factor Model is based on the ideas of a declarative and procedural memory system (Ullman, 2001) as well as the distinction between implicit competence and explicit knowledge (Paradis, 2009); in other words, it assumes that an L2 grammar learned after puberty is stored as a high degree of explicit metalinguistic knowledge in the declarative memory while the native grammar is stored in the procedural memory as implicit competence.

Bardel and Falk (2007) argued that L3 learners, especially adults, accommodate the L3 grammar in the declarative memory alongside the L2 grammar; thus, the L2 grammar is the only accessible source of transfer in learning the L3 (Bardel & Sánchez, 2017; Falk & Bardel, 2010, 2011). In the most recent version of the L2 Status Factor Model, Bardel and Sánchez (2017) incorporated considerations of individual variability in multilingual situations where L1 and L2 were acquired (e.g., simultaneous or sequential bilingual) into the model. They argued that individual differences such as explicit metalinguistic knowledge of L1 or L2, psychotypology, and working memory should be taken into account in L3 acquisition. Among these factors, working memory was particularly highlighted as a critical variable that potentially constrains the access to linguistic representations stored in declarative memory (Baddeley, 2003). Due to its role in suppressing the activation of inappropriate linguistic routines, learners with lower working memory were more likely found to be vulnerable to negative transfer (Fehringer & Fry, 2007; Green, 1998).

The Typological Primacy Model offered by Rothman (2010, 2011, 2013, 2015) is primarily based on the typological proximity among L1, L2, and L3. According to this model, the language parser unconsciously identifies the typological proximity based on the hierarchy of various linguistic cues, including lexicon, phonology/phonotactics, functional morphology, and syntactic structure (Rothman, 2013, 2015; Rothman, González Alonso & Puig-Mayenco, 2019). Once the structurally and typologically closer language is decided by the parser, wholesale transfer occurs through the properties of the selected language at the initial stage of L3 acquisition.

Meanwhile, Fallah, Jabbari and Fazilatfar (2016) suggested the idea of the privileged status of a more dominant language between L1 and L2 in L3 acquisition, with the language that serves communicative functions constituting the default language of transfer irrespective of the order of language acquisition (Fallah & Jabbari, 2018). Although the model does not fully reject the possibility of language transfer as a consequence of the privileged role of L2 or typological proximity between languages, it further suggests the role of dominant language, claiming that any dominant language can be a source of transfer (Jabbari, Achard-Bayle & Ablali, 2018).

Some models do not specify a particular language for transfer. Flynn, Foley and Vinnitskaya (2004) proposed the Cumulative Enhancement Model, claiming that transfer is language-selective and either facilitative or neutral (Berkes & Flynn, 2012; Flynn & Berkes, 2017). Others hold that various factors may play a modulating role in L3 transfer, such as the linguistic properties of interest, learner's age, recency, instruction, and language dominance (Westergaard, 2019).

Previous studies have been predominantly tested with adult learners, with only a few studies investigating the validity of the models in the context of children's L3 acquisition. For example, Hopp (2019) conducted sentence repetition and oral sentence production tasks with L1 Turkish, L2 German child learners of L3 English between 10 and 11 years of age. The main focus of the study was to examine whether the children would transfer the specific word order (verb-second, adverb position, and verb-complement) from their L1 or L2 in their production of L3 sentences. The results showed that the children's L3 production echoed the word order patterns in the L2. Hopp concluded that the results support the Typological Primacy Model because L3 English is typologically more similar to L2 German than to L1 Turkish. He further claimed that the results are irrelevant to the L2 Status Factor Model, which is theoretically based on the declarative and procedural memory system, because the model cannot be applied to children who are assumed to capitalize mostly on the procedural memory system to learn languages.

Recent studies of L3 acquisition have directly pointed out that factors other than language variables, such as language input, use, age, language dominance, and language proficiency, need to be investigated to understand such a complex aspect of L3 acquisition (Hopp, 2019; Slabakova, 2017; Westergaard, 2019). Therefore, examining various factors in relation to language variables offers the key for expanding our understanding of the main interests of L3 acquisition (i.e., L3 initial state and L3 grammar development).

Present study

This study examined the interpretive preference for negation and quantifier scope in the L3 English by two groups of multilingual children with different L1s (Russian and Chinese) but the same L2 (Korean). As in Hopp (2019), this study aims to test the validity of the L3 acquisition models, but we also included three

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

other potential factors – namely, proficiency, language dominance, and working memory – to identify and characterize the source of the transfer underlying L3 acquisition more precisely.

Participants

The study involved 118 children: 50 L1 Korean child learners of L2 English (L2 group, aged 11;5–12;3) and 68 L1 Chinese or Russian child learners of L3 English (L3 group, aged 11;4–12;9). All were recruited from a local elementary school in South Korea. The L3 English children consisted of 34 L1 Mandarin Chinese and 34 L1 Russian. All of them were born in China or Russian-speaking countries (21 in Uzbekistan, 7 in Kazakhstan, and 6 in Russia) and immigrated to Korea with their parents, who took jobs in Korea. We included the Chinese- and Russian-speaking children as the same L3 group because they constituted the two largest populations from immigrant families in the school at the time of testing. Due to their historical and geographical bonds with Korea, the number of immigrant people from China and Russia (and nearby countries, such as Uzbekistan and Kazakhstan) has rapidly increased, forming a multiethnic, multicultural society in Korea (Hong, 2010). Examining children from these two language groups as the representative L3 group may allow for the generalization of our findings to the growing immigrant populations within Korea. In addition, the comparable syntactic patterns across Chinese and Russian in terms of preferences of quantifier scope interpretation, as previously outlined, also justify our inclusion of the two language groups as the homogenous L3 group. The mean length of residence was 3;1 years for the L1 Chinese group (range = 12–60 months) and 2;4 years for the L1 Russian group (range = 10–60 months). All participants attended the same public elementary school in Korea at the time of data collection. After entering the school, they all took an intensive Korean language class for two hours a day and studied general subjects in Korean. According to a teacher at the school, none of the students showed difficulty in communicating in L2 Korean and learning the general subjects in school; they all could read and write Korean without difficulty. Following the regular school curriculum, they started learning English at school from age 9. Although most participants had their first exposure to English after their exposure to Korean, 23 children had already been exposed to English before they came to Korea. However, their exposure to English in their home country was very restricted in the amount of input, mostly confined to a small number of simple words and expressions from the media. In other words, all the participants started learning English through regular curriculum after they had received substantial exposure to Korean.[6] Therefore, following the definition provided by Hammarberg (2010) that an L3 refers to "a non-native language which is currently being used or acquired in a situation where the person already has knowledge of one or more L2s" (p. 97), we determined English as the Chinese- and Russian-speaking children's L3. The English classes were provided for only three hours per week, with a focus on teaching vocabulary or very simple sentences. As in the case of their Korean-speaking peers at the same age, most participants could hardly produce correct sentences in English, as reflected in their poor performance in the picture narration task (described below), suggesting that they were beginners and still in the initial state of L3 English acquisition.

[6]While these 23 children had earlier exposure to English, their English proficiency did not significantly differ from the other participants at the time of data collection ($t(66) = -2.33$, $p = .82$). Moreover, when we reanalyzed the English TVJT results for this subset group only, the analysis showed the same results (full-set reading 92.0% vs. partitioned-set reading 75.9%) as those based on the full datasets (full-set reading 93.6% vs. partitioned-set reading 75.4%, see Figure 3 for details). This suggests that the earlier exposure to English may not have affected their proficiency in this subset group. Based on these observations, we categorized these 23 children as beginner-level L3 English learners and included them in the same group as the other L3 learners.

Materials and procedures

Truth-value judgment task

To investigate the children's interpretive preferences for negation and quantifier scope, this study used a truth-value judgment task (TVJT), following the work in previous studies (O'Grady, 2013; O'Grady et al., 2009a; O'Grady et al., 2011). The task included six short stories accompanied with pictures presented on a laptop computer. Considering the children's low proficiency in English, the stories were presented in both the L1 (Chinese or Russian) and the L2 (Korean), which were written on each slide; however, the test sentences (e.g., *Tom did not cut all the trees*) were presented in their L3 English. The task was conducted individually in a quiet room, and each child was asked to judge whether the test sentence corresponded to each story. There were three test items for each of the two conditions (full-set interpretation and partitioned-set interpretation), and the items were arranged in a Latin square design so that each participant encountered only one condition of a single item. The sample stories supporting the full-set and partitioned-set interpretations are exemplified in (2) and (3).

(2) Sample story in the full-set condition
Today, Tom made up his mind to clean six windows and cut down three trees. Tom cleaned the six windows right away. Next, Tom tried to cut down the first tree, but it was too tall to cut down. Then, Tom tried to cut down the next tree, but it was also too tall to cut down. After that, Tom tried to cut down the third tree. Again, it was too tall to cut down, so he decided not to do it.
Test sentence: Tom did not cut all the trees.

(3) Sample story in the partitioned-set condition
Today, Tom made up his mind to clean six windows and cut down three trees. Tom cleaned the six windows right away. Next, Tom cut down the first tree, and it was very easy. Then, Tom cut down the second tree. It was a bit harder, but it was still fine. After that, Tom tried to cut down the third tree. It was too tall to cut down, so he decided not to do it.
Test sentence: Tom did not cut all the trees.

The experiment also included four practice and four filler items. The practice items, presented at the beginning of the task, contained clearly unambiguous sentences like *This is Jerry* and *Jerry is not a cat* paired with a picture showing a mouse named *Jerry*. These items were included to ensure that the children familiarized themselves with the task and to test whether they could properly match the statement with the corresponding picture. The filler items were presented in the same manner; we used unambiguous sentences like *Jerry did not play the violin* paired with a picture showing Jerry playing the piano. Half of the practice and filler items showed a congruity between the statement and the picture, eliciting a "yes" response; the other half showed a mismatch between the statement and the picture,

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

eliciting a "no" response. The target answers for the practice and filler items were counterbalanced.

After each child completed the experiment, the researcher asked him/her whether there had been any English words or sentences that he/she did not know the meaning of, and all of them replied that they knew the meaning of the words and sentences. The Chinese- and Russian-speaking children were individually tested first in L3 English. Two weeks after the L3 experiment, they completed the same tasks with L1 or L2 materials in a paper-and-pencil manner. Two weeks later, they took the task again in their other language. Half of the children completed the L1 task before the L2 task, and the other half completed it after. The Korean-speaking children in the control group completed the English task before they completed the Korean version of the task two weeks later.

Picture narration task

Because language proficiency has been considered a crucial factor, as revealed in previous L2 acquisition studies (e.g., Song & Schwartz, 2009) as well as in L3 acquisition studies (e.g., Arıbaş & Cele, 2019; Jaensch, 2012; Sánchez & Bardel, 2017), it is important to measure learners' proficiency in each language. To this end, a picture narration task in each language version was employed (Song & Schwartz, 2009). The task consists of three sets of four pictures that describe a sequence of daily events, such as washing one's face, eating food, and reading a book. The series of pictures were presented in Microsoft PowerPoint, and each child was prompted to provide oral descriptions of the events in the respective languages. The participants' descriptions of each of the 12 pictures were recorded, transcribed, and scored based on the complexity and accuracy of the speech (see Song & Schwartz, 2009 for details on the assessment procedure).

Assessment of language dominance: The HALA project

To test the potential role of language dominance in L3 acquisition (Fallah et al., 2016), we adopted the Hawaii Assessment of Language Access (HALA) task to determine the children's language dominance (L1 or L2). The HALA project, developed by O'Grady, Schafer, Perla, Lee and Wieting (2009), exploits a picture-naming task to measure the speed of bilingual speakers' access to words for body parts in their two languages. Its theoretical foundation is grounded on the well-attested relationship between word frequency and accessibility – namely, that the increased use of a language leads to improved lexical accessibility. The body-part naming task consists of 31 test items. In the current study, the children were tested in their L1 and L2 at an interval of at least two weeks. We determined the relative strength of language dominance across the two languages by considering both reaction times (RTs), operationalized as the duration between the onset of the picture on the screen and the onset of the response from a participant in milliseconds, as well as the response accuracy.[7]

[7]We prioritized RT results over the accuracy when determining the dominant language according to the assessment procedure employed by the test developers (O'Grady et al., 2009a). When the participant's RT difference between the L1 and L2 tasks was within 100 milliseconds (10 cases in our study), the dominant language was decided in reference to the average accuracy score (for details of the data coding procedures, see Hamilton, Perla & Robinson, 2013).

Working memory task

The most recent version of the L2 Status Factor Model explains individuals' working memory capacity (WMC) as a primary source of determining the source of transfer (Bardel & Sánchez, 2017). To test the validity of this model, we measured the participants' WMC using an operation span (OSPAN) task. We opted for the OSPAN task as the measure of the children's WMC because it is a nonverbal task and thus suitable for assessing WMC of beginner-level learners, who have difficulty processing linguistic materials (for details on an OSPAN task, see Linck, Osthus, Koeth & Bunting, 2014; Wen, 2016). The task was also appropriate for our objective of measuring domain-general cognitive abilities from our participants who consisted of learners with different L1 backgrounds. We used the computerized version of the OSPAN task provided by PEBL (Mueller & Piper, 2014). The OSPAN task presented a total of 42 mathematical operations (e.g., 9 + 2–3 = ?), and the participants decided whether the answer presented following each operation (e.g., 8) was correct or not. After the participants verified the answer to the arithmetic operation, they were presented with strings of letters from the English alphabet as recall items. The mathematical operations were distributed in sets of 2, 3, 4, and 5, with three sets for each set size, and the set sizes were randomly presented. PEBL automatically recorded the RTs and the correctness of the answers to the mathematical operations and the alphabet letter recall.

Results

Proficiency, language dominance, and working memory capacity

Table 1 provides the summary of the children's L2 proficiency, dominant language, and WMC between the two language groups.[8]

Of the 68 children total, 19 Chinese children and 5 Russian children belonged to the high proficiency group, 10 Chinese and 14 Russian to the mid-proficiency group, and 5 Chinese and 15 Russian to the low-proficiency group. Participants' L2 proficiency was assessed by calculating both syntactic complexity (i.e., the total number of words divided by the total number of T-Units) and accuracy (i.e., the rate of error-free T-Units) scores. We converted the scores from the two indices into z-scores and averaged them to obtain a composite proficiency score (for details of the procedures used to measure proficiency, see Song & Schwartz, 2009). Although lengths of residence were similar between the two groups, L2 Korean proficiency was higher in the L1 Chinese group ($M = 14.1$, $SD = 5.7$) than in the L1 Russian group ($M = 11.4$, $SD = 5.3$) ($t(66) = 4.43$, $p < .001$, Cohen's $d = 0.515$).[9]

As for language dominance, as measured by the HALA task, 51 children showed L1 dominance, showing faster RTs and

[8]An important issue related to WMC concerns scoring methods. As Leeser and Sunderman (2016) pointed out, there is no agreement on the best scoring method, and different methods can lead to different results in individuals' WMC scores. This study applied all five scoring methods (i.e., recall, recall + accuracy, recall + accuracy + RT, set size, and composite Z score) that Leeser and Sunderman used and compared. We found no difference in WMC between the two multilingual groups, regardless of the scoring method used. The WMC score presented in Table 1 is a composite Z score, which treated all related scores together.

[9]A potential explanation for this asymmetry is typological proximity. Korean is typologically closer to Chinese than Russian, which may facilitate Chinese learners' acquisition of Korean. In addition, a lot of Korean words come from Chinese, which again helps them learn Korean.

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

Table 1. Participant information on L2 proficiency, dominant language and WMC

		L2 proficiency			Dominant language		WMC	
Group	N	High	Mid	Low	L1	L2	M	SD
L1 Chinese	34	19	10	5	18	16	0.02	0.7
L1 Russian	34	5	14	15	32	2	−0.03	0.56

Table 2. Results of the HALA tasks for participants

	L1		L2			
	Mean	SD	Mean	SD	t	*p*
HALA: Accuracy (%)	88.0	14.8	68.6	21.8	5.397	<.001
HALA: Response time (ms)	1305.6	249.1	1418.1	267.3	−3.135	.003

more accurate results in the L1 over the L2, while the other 17 children showed L2 dominance. Table 2 presents participants' mean accuracy and response time in the HALA tasks.

To compare participants' language dominance across the L1 and L2, paired-samples *t*-tests were conducted on the accuracy and RTs in the HALA tasks and self-reported language use. Results showed significant differences between the L1 and L2 in terms of both accuracy ($t(67) = 5.397$, $p < .001$, Cohen's $d = 1.040$) and RTs ($t(67) = -3.135$, $p = .003$, Cohen's $d = 0.435$) in the HALA tasks as well as in their self-reported language use ($t(67) = 3.464$, $p = .001$, Cohen's $d = 0.687$), indicating that the learners were more accurate and faster in their responses in the L1 than L2 HALA task and that they used L1 more often than L2 in daily life. These results confirm an overall dominance of L1 rather than L2 in the children.

Truth-value judgment task in L1

We analyzed participants' responses in the task in terms of preferences for a full-set or partitioned-set interpretation for each item. As displayed in Figure 1, both L1-Chinese and L1-Russian learners showed strong preferences for both full-set and partitioned-set interpretations, accepting each condition more than 90% of the time. In each language group, the L1-Chinese group accepted a full-set interpretation ($M = 98.0\%$, $SD = 13.9\%$) more often than a partitioned-set interpretation ($M = 92.2\%$, $SD = 27.0\%$). Similarly, the L1-Russian group accepted a full-set interpretation ($M = 97.1\%$, $SD = 17.0\%$) more often than a partitioned-set interpretation ($M = 95.1\%$, $SD = 21.7\%$). Despite the numerical trend toward more acceptance for a full-set than for partitioned-set interpretation in both groups, the acceptance difference between the conditions was very small, indicating the children's overall acceptance for both conditions.

The proportion of preferences for the two conditions in each group was statistically compared using a logistic mixed-effects regression (Baayen, 2008). The model was created in R (R Core Team, 2014) using the *glmer* program in the *lme4* package and included participants' L1 (Chinese, Russian), condition (full-set interpretation, partitioned-set interpretation), and interaction as fixed effects, along with the random effects of participant and item. We initially constructed the maximal random effects structure allowed by the design (Barr, Levy, Scheepers & Tily, 2013) and then simplified the structure from more complex to less complex in a progressive way through log-likelihood ratio comparisons to reduce a potential loss of statistical power (Matuschek, Kliegl, Vasishth, Baayen & Bates, 2017). All fixed factors were contrast-coded and centered around the mean.

The model showed no main effect of L1 ($\beta = -0.128$, $SE = 3.087$, $p = .967$), no main effect of condition ($\beta = -0.198$, $SE = 2.388$, $p = .934$), and no interaction between the two factors ($\beta = 0.752$, $SE = 4.614$, $p = .870$), indicating that both L1-Chinese and L1-Russian learners accepted the full-set and partitioned-set conditions to the same extent. The results of the L1 TVJT confirmed that both full-set and partitioned-set interpretations are allowed in the L1 grammars of the children. By virtue of the consistent patterns between the two language groups in the L1 task, we collapsed the L1 factor in the subsequent analyses, combining both L1-Chinese and L1-Russian children as a single group.

Truth-value judgment task in L2 Korean

As in the previous analysis, we inspected participants' responses in the L2 task in terms of preferences for a full-set or partitioned-set interpretation. The L2 group's performance on the task was compared to the control group of L1-Korean children. As shown in Figure 2, the L1-Korean and L2-Korean groups exhibited a similar pattern in the proportion of acceptance for the full-set and partitioned-set conditions. The L1-Korean group was more likely to accept the full-set ($M = 84.0\%$, $SD = 16.0\%$) than the partitioned-set condition ($M = 23.5\%$, $SD = 16.0\%$). Likewise, the L2-Korean group accepted the full-set condition ($M = 91.7\%$, $SD = 18.9\%$) more often than the partitioned-set condition ($M = 56.4\%$, $SD = 18.9\%$), although their acceptance rates for the partitioned-set condition were greater than those in the L1-Korean group.

To statistically compare the group performance across the conditions, we ran logistic mixed-effects regression (*glmer*) in the same manner as in the previous analysis. The model included group (L1-Korean, L2-Korean), condition (full-set interpretation, partitioned-set interpretation), and their interaction as fixed effects (contrast-coded and centered) and the random effects of participant and item. Again, the random effects structure was progressively reduced from the maximal to a simpler structure via

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

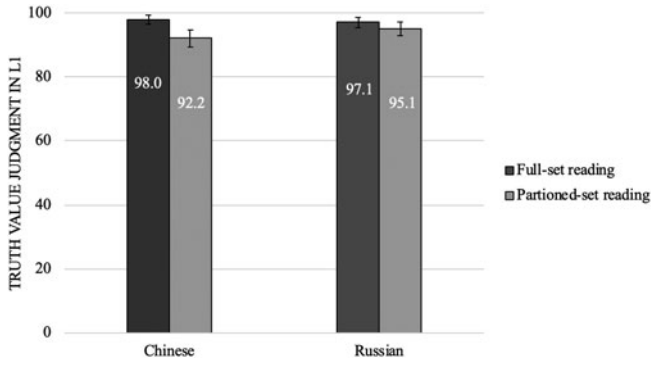

Fig. 1. Mean percentage of acceptance in the L1 TVJT (Chinese and Russian); error bars denote 95% CIs

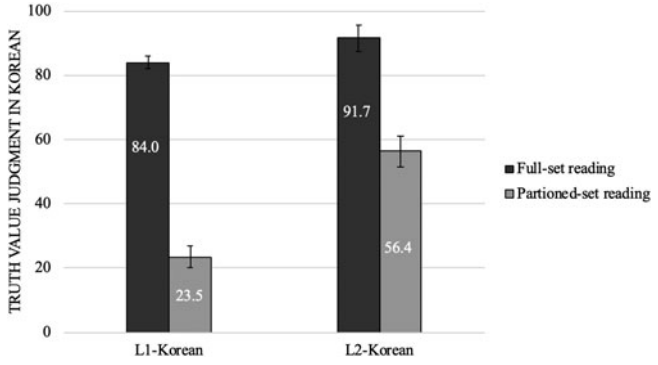

Fig. 2. Mean percentage of acceptance in the Korean TVJT; error bars denote 95% CIs

Table 3. Logistic mixed-effects results in the Korean Truth-Value Judgment Task

	ß	*SE*	*p*
(Intercept)	1.060	0.380	.005
Group	−1.686	0.442	<.001
Condition	−2.675	0.725	<.001
Group × Condition	−1.780	0.751	.018

log-likelihood ratio tests. The model outcomes are summarized in Table 3.

The model revealed a main effect of group, with higher acceptance rates in the L2-Korean than the L1-Korean group, and a main effect of condition, induced by higher acceptance rates for the full-set than the partitioned-set interpretation. These main effects were qualified by an interaction of the two factors, indicating that the difference of acceptance rates between the conditions differed across the groups. In light of this interaction, we built separate models for each group, which included the condition as a fixed effect with the random effects of participant and item. The by-group analyses demonstrated a main effect of the condition in both the L1-Korean group ($\beta = -3.706$, $SE = 0.817$, $p < .001$) and the L2-Korean group ($\beta = -2.188$, $SE = 0.760$, $p = .004$). These results indicate that both groups had a strong preference for the full-set interpretation in the Korean TVJT, although the difference was greater in the L1-Korean than in the L2-Korean group.

To further explore the effect of proficiency in the L2 data, we added participants' scores on the Korean picture narration task to the L2 model as a continuous variable. The model again showed a main effect of condition ($\beta = -2.434$, $SE = 0.790$, $p = .002$), consistent with the results of the by-group analyses. Importantly, there was a significant interaction between condition and L2 proficiency ($\beta = -0.197$, $SE = 0.071$, $p = .005$), indicating that the L2 group had a stronger tendency to accept the full-set than partitioned-set condition as their L2 proficiency was higher. Given that the L1-Korean group showed a greater gap in the acceptance rates between the two conditions than the L2-Korean group, this interaction in the L2 group suggests their convergence on target-like interpretations with increasing L2 proficiency.

In summary, despite an overall dominance toward the L1 and the tendency of equally accepting both conditions in the L1 TVJT, the L2-Korean group showed interpretations similar to those of the native speaker group in the Korean TVJT. Moreover, the learners were more likely to display target-like performance as their Korean proficiency increased. The L2-Korean group's distinct patterns across the L1 and L2 tasks offered an ideal case to test which language between the learners' L1 or L2 played a more privileged role for transfer in the distributive scope interpretations in L3 English, the main question addressed in this study.

Truth-value judgment task in L3 English

In this task, the L1-Chinese and L1-Russian children were included as an experimental group (L2-Korean group, $n = 68$), while another group of Korean-speaking children, who did not participate in the L2 task, served as a comparison group (L1-Korean group, $n = 23$).

Before analyzing the results of the L3 English TVJT, we inspected participants' performance on filler items and removed as outliers those who scored less than 75% of the time (i.e., more than 2 incorrect answers in total for the four practice and four filler items). This was done to ensure that participants had a basic understanding of the English sentences they encountered during the task so that their performance on the English task would be meaningfully interpreted. This process resulted in a loss of data from 11 participants in the L2-Korean group, leaving 57. The L1-Korean children all scored higher than 75% on the filler items and, thus, all 23 participants in this group were included in the analysis.

For the remaining participants, the proportion of acceptance for the full-set and partitioned-set conditions in the English task was analyzed across the groups. As displayed in Figure 3, both L1-Korean and L2-Korean groups accepted the full-set condition (L1-Korean: $M = 100\%$, $SD = 0$; L2-Korean: $M = 93.6\%$, $SD = 18.9\%$) more often than the partitioned-set condition (L1-Korean: $M = 52.2\%$, $SD = 21.8\%$; L2-Korean: $M = 75.4\%$, $SD = 17.0\%$).

Participants' performance on the English TVJT was analyzed in detail using a logistic mixed-effects regression in the same manner as in the previous analyses. As summarized in Table 4, the model showed a main effect of condition, with higher acceptance rates for the full-set than the partitioned-set condition, but without a main effect of group or an interaction. These results indicate that both groups had a stronger preference for the full-set versus partitioned-set interpretation in the English TVJT, mirroring their strong preference for the full-set interpretation in the

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

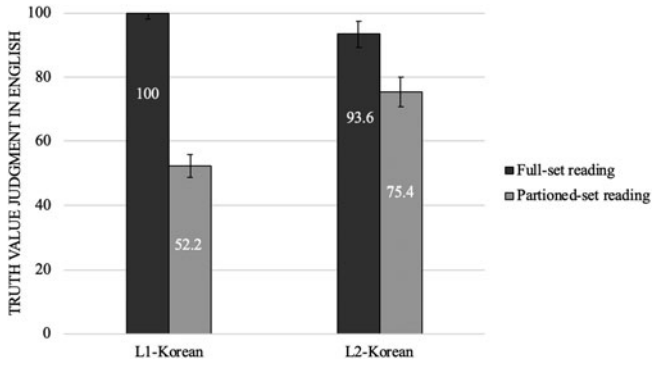

Fig. 3. Mean percentage of acceptance in the English TVJT; error bars denote 95% CIs

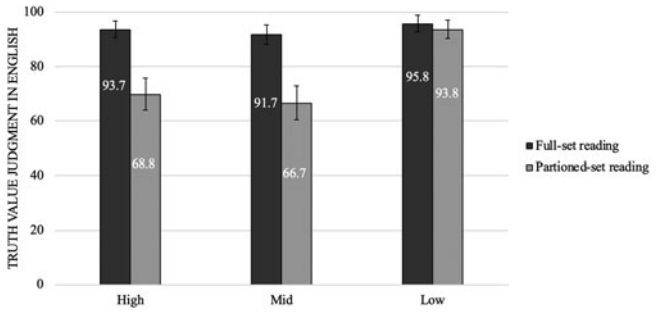

Fig. 4. Mean percentage of acceptance in the English TVJT by L2 proficiency group; error bars denote 95% CIs

Table 4. Logistic mixed-effects results in the English Truth-Value Judgment Task

	ß	*SE*	*p*
(Intercept)	4.567	0.930	<.001
Group	−0.487	1.152	.673
Condition	−5.472	1.756	.002
Group × Condition	−3.795	2.320	.102

Korean TVJT. The by-group analyses further confirmed the main effect of condition in both the L1-Korean group ($\beta = -6.260$, $SE = 1.678$ $p < .001$) and L2-Korean group ($\beta = -1.898$, $SE = 0.412$ $p < .001$), indicating that both groups had a strong preference for the full-set interpretation in the English TVJT.

Next, we investigated the potential effects of the learners' language dominance, working memory capacity, L1 proficiency, and L2 proficiency in their L3 English scope interpretations. To this end, we created separate mixed-effects models for the L2-Korean group, including participants' scores on the working memory task, the language dominance (L1 dominant, L2 dominant), the scores on the L1 picture narration task, and the scores on the L2 Korean picture narration task, respectively, as an additional interactive factor.

When the model included condition, language dominance, and their interaction as fixed effects, there was only a main effect of condition ($\beta = -1.904$, $SE = 0.418$, $p < .001$), without an effect of language dominance ($\beta = -0.007$, $SE = 0.044$, $p = .878$) or its interaction with condition ($\beta = -0.081$, $SE = 0.061$, $p = .181$). The model including working memory capacity as an additional factor showed a main effect of condition ($\beta = -2.087$, $SE = 0.465$, $p < .001$) and a main effect of working memory capacity ($\beta = 0.984$, $SE = 0.492$, $p = .045$), yet there was no interaction between the two factors ($\beta = -0.872$, $SE = 0.695$, $p = .210$). These results indicate that the participants' language dominance or their working memory capacity had little impact on their distinguished scope interpretations between the conditions in the English task.

Similarly, the model including condition, L1 proficiency, and their interaction as fixed effects showed only the main effect of condition ($\beta = -1.907$, $SE = 0.414$, $p < .001$), without an effect of L1 proficiency ($\beta = 0.033$, $SE = 0.070$, $p = .637$) or the interaction between condition and L1 proficiency ($\beta = -0.024$, $SE = 0.091$, $p = .790$). These results indicate that the L2-Korean participants had greater acceptance rates for the full-set condition than for the partitioned-set condition, regardless of their proficiency in the L1.

Next, we included participants' scores on the L2 Korean picture narration task in the model. The model showed a robust effect of condition ($\beta = -1.860$, $SE = 0.419$, $p < .001$), induced by their stronger preference for the full-set than partitioned-set interpretation, and there was no main effect of L2 proficiency ($\beta = -0.034$, $SE = 0.053$, $p = .524$). Notably, there was an approaching significance in the interaction between condition and L2 proficiency ($\beta = -0.139$, $SE = 0.075$, $p = .065$), suggesting that the acceptance difference between the full-set and partitioned-set conditions was larger as participants' L2 proficiency was higher. In light of this marginal interaction, and to fully explore the role of L2 proficiency, we divided the L2-Korean participants into three groups based on their Korean proficiency (21 in high, 20 in mid, 16 in low group) and conducted separate analyses for each proficiency group.

Figure 4 illustrates the results of the English TVJT by L2 proficiency groups. The three groups demonstrated distinct patterns: While the high and mid groups accepted the full-set more than the partitioned-set condition, the low group accepted both conditions to a similar extent.

Separate models for each proficiency group were created, including condition as a fixed effect (contrast-coded and centered) and participant and item as random effects. The random-effects structure was simplified in the same manner as in the previous analyses. The model for the high group showed a main effect of condition ($\beta = -2.326$, $SE = 0.965$, $p < .001$), with higher acceptance rates for the full-set than the partitioned-set condition. Similarly, a robust effect of condition was found in the model for the mid group ($\beta = -2.003$, $SE = 0.622$, $p = .001$). However, the effect of condition was not significant in the model for the low group ($\beta = -0.477$, $SE = 0.978$, $p = .626$), indicating their acceptance of both conditions.

Taken all together, the analyses of the English TVJT showed that the L1-Korean and L2-Korean groups were comparable in their scope interpretations of the English sentences, with a stronger preference for the full-set than for the partitioned-set interpretation. Considering that the bias toward a full-set interpretation was found only in the Korean TVJT, but not in the Chinese and Russian tasks, the performance of the L2-Korean children in the English TVJT provides compelling evidence of their transfer of L2 properties. Furthermore, there was no interacting role of the participants' language dominance or working memory capacity in their task performance, yet the degree of

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

transfer from the L2 was augmented as learners' L2 proficiency increased.

Discussion

The major findings of this study can be summarized as follows. First, the children showed a transfer of L2 properties in their interpretation of negation and quantifier scope in L3 English. Second, neither language dominance nor WMC affected the children's interpretation of L3 English. Third, the learners with higher L2 proficiency showed a higher degree of crosslinguistic influence from the L2 to the L3. Based on these findings, this section evaluates the models of L3 acquisition.

The Cumulative Enhancement Model proposes that prior language experience can either enhance subsequent language acquisition or have no effect. This model is incompatible with our findings as it cannot explain the children's rejection of the partitioned-set interpretation in the L3 English sentences, which shows the negative transfer of L2 Korean. The model predicting transfer from L1 also fails to explain the current results because both the L1-Chinese and L1-Russian children carried over the L2 Korean, but not their L1, interpretive tendency.

At a first glance, the results seem to support the L2 Status Factor Model, which proposes a crucial role of L2 transfer in L3 acquisition, as our L2-Korean children showed the same patterns across the L2 and L3 tasks. However, as Hopp (2019) noted, this model cannot apply to children's L3 acquisition because the model focuses on learners learning an L3 after puberty. As previously reviewed, this model assumes that both L2 and L3 acquisition are mainly driven by the declarative memory system, whereas L1 acquisition is assisted by the procedural memory system. This idea is difficult to reconcile with children's L3 acquisition in which all three languages are assumed to be acquired through the same learning procedure – namely, the procedural memory system.

Further support for the claim that the L2 Status Factor Model cannot explain children's L3 acquisition comes from the incompatibility of our findings with two specific predictions that follow from the model. For example, Sánchez and Bardel (2017) discussed the role of L2 proficiency in L3 acquisition. Assuming that language transfer in L3 acquisition is caused by explicit metalinguistic knowledge of the L2 stored in the declarative memory system (Paradis, 2008) and that the L2 knowledge of adult L3 learners with low L2 proficiency is likely characterized as explicit metalinguistic knowledge, Sánchez and Bardel proposed a distinct effect of L2 transfer, particularly for learners with low L2 proficiency. Our results, however, are in sharp contrast to this explanation: the learners with high and intermediate L2 Korean proficiency – but not the learners with low L2 proficiency – transferred their L2 interpretive preference to the interpretation of the L3 English sentences. In addition, Bardel and Sánchez (2017) discussed the role of working memory for the L2 Status Factor Model. In tandem with the idea of language transfer from explicit metalinguistic knowledge, they highlighted the effect of WMC, claiming that L2 learners, particularly low working memory learners, are likely to employ linguistic knowledge from L2 and thus show more L2-like performance in L3. However, our study did not find any interaction of WMC in L3 English scope interpretation; the children transferred the L2 properties, regardless of their cognitive ability.

The role of dominant language also fails to predict the results of this study. Jabbari et al. (2018) argued that language items in a dominant language are more accessible and thus more likely transferred. Contrary to the model's prediction that language dominance affects L3 acquisition, the mixed-effects regression model did not find any meaningful interaction between the interpretation of the two conditions and language dominance as measured by the HALA task.

Some may find our results supportive of the Linguistic Proximity Model (Westergaard, Mitrofanova, Mykhaylyk & Rodina, 2017) and the Scalpel Model (Slabakova, 2017) because the children in this study showed both facilitative and non-facilitative influence from both previously acquired languages. That is, they rejected the partitioned-set interpretation as the L1-Korean control group did, but not to the same extent. However, as Puig-Mayenco, González Alonso and Rothman (2020) pointed out, the selective and property-by-property transfer proposed by the Linguistic Proximity Model and the Scalpel Model neither specifies the mechanisms underlying the evaluation of the structural similarity nor elucidates different degrees of weights among the multitude of factors involved in determining the source of transfer. Moreover, the theoretical assumption of property-by-property of transfer (Westergaard, 2019) is not compatible with both a cognitive economy perspective (Rothman et al., 2019) and a generative approach associated with pursuing the function of a complete grammar (Schwartz & Sprouse, 2019). In addition, the models are unable to make any specific predictions, especially when non-facilitative transfer is predicted (Rothman et al., 2019). In conclusion, as the models with property-by-property transfer lack a solid theoretical rationale, we argue that the results of the study are not compatible with the models, although they may offer a flexible explanation for transfer in L3 acquisition.

Finally, our results may or may not be compatible with the prediction of the Typological Primacy Model (TPM) depending on the linguistic level(s) at which the parser determined proximity between the languages. This model posits that the initial state of L3 grammar is constrained by what the parser takes to be structurally similar among the L1–L2–L3 grammars based on the similarity in terms of the structure of the lexicon, phonology/phonotactics, functional morphology, and syntactic structure (Rothman, 2015). Notably, the TPM would predict the L1-Russian children's reliance on the knowledge of L1 Russian – not L2 Korean – in the comprehension of L3 English at the lexical and phonological levels because English shares a greater degree of similarity with Russian than with Korean in terms of lexical and phonological properties, the first two steps of the four-level hierarchy. Note that a substantial number of cognates exist between Russian and English (Elgort, 2013; Laufer & McLean, 2016), but only a limited number of cognates exist between Korean and English (Kim & Davis, 2003); in addition, both English and Russian are stress-timed languages whereas Korean is syllabled-timed (Arvaniti, 2012; Nespor, Shukla & Mehler, 2011). Despite the closer proximity of English to Russian than to Korean in the lexicon and phonology levels, our Russian-speaking children showed L2 Korean transfer in the interpretation of L3 English sentences, contrary to the prediction of the TPM. However, we may not preclude the possibility that the children failed to notice the lexical and phonological similarities between the languages and moved to the morphological level to check linguistic proximity. At the morphological level, English is presumably considered more similar to Korean than to Russian because Russian has a distinctively rich inflectional system compared to Korean and English. Russian uses verbal inflection to mark tense, mood, number,

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

gender, and person (Pesetsky, 1979). Korean and English verbal inflection also marks tense and number, but their functions can be relatively limited compared to the Russian inflectional system. If the parser of the Russian-speaking children focused on the morphological similarities between English and Korean, their L2 Korean transfer might be explained by the TPM.

The TPM might also explain the data from L1-Chinese children in some ways. At the lexical level, the parser would find it difficult to make choices about proximity between the languages because English is dissimilar to both Korean and Chinese in terms of the lexicon, although Chinese and Korean have much lexical information in common (Sohn, 1999). At the phonological level, the parser would make no decision either, because English is stress-timed, Korean is syllable-timed, and Chinese has both features. However, at the morphological level, the parser would choose Korean. Both English and Korean have substantially more inflectional and derivational morphemes than Chinese, an isolating language. Thus, it is possible that the Chinese speakers attended to the similarities between the languages at the morphological level and chose L2 Korean as the primary source of transfer in their L3 English interpretation.

Ultimately, we conclude that no current model of L3 acquisition can successfully explain this study's results, except for the TPM. The TPM might account for the observed data if we assume that the Russian- and Chinese-speaking children noticed the interlanguage similarities at the morphological, but not the lexical and phonological levels. However, we have no clear theoretical motivation for the possibility that the two L1 groups would make decisions on linguistic proximity at only the morphological level. To provide the most likely and unified explanation of the current findings, we focused on the role of community language proficiency in L3 acquisition. It is widely attested that the speakers' proficiency in their community language is strongly associated with the amount of exposure to the language. For example, previous heritage language acquisition studies show that heritage speakers are usually more exposed to their community language (or L2) than the heritage language (or L1), thereby evincing higher proficiency in the community language than their heritage language and even L1 attrition (Montrul, 2010). As the parser is "tuned to the frequency of particular structures in the language the hearer hears" (Carroll, 2001, p. 10), it probably looks to the community language as the source of transfer because of more experiences in the community language than their L1. Previous studies on L3 acquisition provide supporting evidence for the role of community language proficiency. For example, Hopp (2019) showed the transfer of L2 German (the community language) to L3 English among participants whose L2 proficiency was high. Although he concluded that the results support the TPM, the results are compatible with the community proficiency account. As in the case of heritage speakers, the children in our study used L2 Korean (i.e., the community language) in a variety of contexts and more often than their L1 (i.e., the heritage language) as their proficiency increased. This indicates that they must have experienced less difficulty accessing and activating the grammar of L2 Korean compared to the lower L2 proficiency children.

In conclusion, our results showed the influence of L2 Korean knowledge for intermediate and high proficiency Korean learners, but not for low proficiency Korean learners. Despite their dominance in the L1, the more proficient groups may have had more experience using their L2 Korean, the community language, than their L1 and thus had more opportunities to activate L2 than L1. Given this possibility, we can speculate that the children with higher L2 Korean proficiency might have been more likely to activate their L2 knowledge in the L3 task, which resulted in the transfer of L2 in their interpretation of L3 English sentences.

Conclusion

This study has shown that multilingual children learning Korean as an L2 and English as an L3 transfer the properties of L2 in their scope interpretations in L3 English. Although several studies and models have been proposed to identify primary sources of transfer in L3 acquisition, none of them appears to fully explain the findings in our study, which points to the need for a new account. Based on this study's results, we claim that, in a context where the L2 is the community language, L2 proficiency plays an important role in the interpretation of the L3 because high-proficiency L2 learners in such a context are likely to have more opportunities to activate their L2 knowledge. We hope that this 'community language proficiency' account will provide numerous opportunities for further studies to test its validity in diverse L3 learning contexts.

Acknowledgements. We would like to thank Hee-Don Ahn, Dong-hwan Lee, Chang-Yong Sim, all our participants, and anonymous reviewers for their very helpful comments and discussions. Some findings in this study have been reported in the first author's unpublished doctoral dissertation. All remaining errors are our own. Correspondence concerning this article should be addressed to the corresponding authors, Kitaek Kim (kitaek@snu.ac.kr) or Hyunwoo Kim (hyunwoo2@yonsei.ac.kr).

References

Arıbaş DŞ and Cele F (2019) Acquisition of articles in L2 and L3 English: the influence of L2 proficiency on positive transfer from L2 to L3. *Journal of Multilingual and Multicultural Development*, doi: 10.1080/01434632.2019.1667364. Published online by Taylor & Francis, September 19, 2019.

Anderson C (2004) The structure and real-time comprehension of quantifier scope ambiguity. Ph.D. dissertation, Northwestern University.

Arvaniti A (2012) The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* **40**, 351–373.

Baayen RH (2008) *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Baddeley A (2003) Working memory and language: An overview. *Journal of Communication Disorders* **36**, 189–208.

Bardel C and Falk Y (2007) The role of the second language in third language acquisition: the case of Germanic syntax. *Second Language Research* **23**, 459–484.

Bardel C and Sánchez L (2017) The L2 status factor hypothesis revisited: The role of metalinguistic knowledge, working memory, attention and noticing in third language learning. In Angelovska T and Hahn A (Eds.), *L3 syntactic transfer: Models, new developments and implications*, pp. 85–101. Amsterdam: John Benjamins.

Barr DJ, Levy R, Scheepers C and Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language* **68**, 255–278.

Berkes É and Flynn S (2012) Further evidence in support of the Cumulative-Enhancement Model. In Cabrelli Amaro J, Flynn S and Rothman J (Eds.), *Third language acquisition in adulthood*, pp. 143–164. Amsterdam: John Benjamins.

Carroll SE (2001) *Input and evidence: The raw material of second language acquisition*. Philadelphia: John Benjamins.

Chung ESE (2013) Sources of difficulty in L2 scope judgments. *Second Language Research* **29**, 285–310.

Elgort I (2013) Effects of L1 definitions and cognate status of test items on the Vocabulary Size Test. *Language Testing* **30**, 253–272.

https://doi.org/10.1017/S1366728920000395 Published online by Cambridge University Press

Green D (1998) Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* **1**, 67–81.
Grice P (1989) *Studies in the way of words*. Cambridge, MA: Harvard University Press.
Falk Y and Bardel C (2010) The study of the role of the background languages in third language acquisition. The state of the art. *IRAL - International Review of Applied Linguistics in Language Teaching* **48**, 185–219.
Falk Y and Bardel C (2011) Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research* **27**, 59–82.
Fallah N and Jabbari AA (2018) L3 acquisition of English attributive adjectives Dominant language of communication matters for syntactic cross-linguistic influence. *Linguistic Approaches to Bilingualism* **8**, 193–216.
Fallah N, Jabbari AA and Fazilatfar AM (2016) Source(s) of syntactic cross-linguistic influence (CLI): The case of L3 acquisition of English possessives by Mazandarani-Persian bilinguals. *Second Language Research* **32**, 225–245.
Fehringer C and Fry C (2007) Hesitation phenomena in the language production of bilingual speakers: The role of working memory. *Folia Linguistica* **41**, 37–72.
Flynn S and Berkes É (2017) Toward a new understanding of syntactic CLI. L3 Syntactic Transfer: Models, new developments and implications. In Angelovska T and Hahn A (Eds.), *L3 syntactic transfer: Models, new developments and implications*, pp. 35–61. Amsterdam: John Benjamins.
Flynn S, Foley C and Vinnitskaya I (2004) The cumulative-enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism* **1**, 3–16.
Hamilton A, Perla J and Robinson LC (2013) A psycholinguistic assessment of language change in eastern Indonesia: evidence from the HALA project. In Jones MC and Ogilvie S (Eds.), *Keeping languages alive: Documentation, pedagogy and revitalization*. pp. 16–28. Cambridge: Cambridge University Press.
Han C, Lidz J and Musolino J (2007) V-Raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry* **38**, 1–47.
Hammarberg B (2010) The languages of the multilingual: Some conceptual and terminological issues. *IRAL-International Review of Applied Linguistics in Language Teaching* **48**, 91–104.
Hermas A (2010) Language acquisition as computational resetting: verb movement in L3 initial state. *International Journal of Multilingualism* **7**, 343–362.
Hermas A (2014) Restrictive relatives in L3 English : L1 transfer and ultimate attainment convergence. *Australian Journal of Linguistics* **34**, 361–387.
Hermas A (2015) The categorization of the relative complementizer phrase in third-language English: A feature re-assembly account. *International Journal of Bilingualism* **19**, 587–607.
Hermas A (2018) Discourse-linking in advanced L3 English: Testing the Interface Hypothesis. *Lingua* **212**, 20–29.
Hong WP (2010) Multicultural education in Korea: Its development, remaining issues, and global implications. *Asia Pacific Education Review* **11**, 387–395.
Hopp H (2019) Cross-linguistic influence in the child third language acquisition of grammar: Sentence comprehension and production among Turkish-German and German learners of English. *International Journal of Bilingualism* **23**, 567–583.
Ionin T, Luchkina T and Stoops A (2014) Quantifier scope and scrambling in the second language acquisition of Russian. In Minai U *et al.* (Eds.), *Selected proceedings of the 5th Conference on Generative Approaches to Language Acquisition–North America*, pp. 169–180. Somerville, MA: Cascadilla Proceedings Project.
Jabbari AA, Achard-Bayel G and Ablali D (2018) Acquisition of L3 French wh-question structure by Persian-English bilinguals. *Cogent Education*, doi: 10.1080/2331186x.2018.1524551. Published online by Taylor & Francis, October 08, 2018.
Jaensch C (2012) L3 Acquisition of German: Do some learners have it easier? In Cabrelli Amaro J, Flynn S and Rothman J (Eds). *Third language acquisition in adulthood*. pp. 165–193. Amsterdam: John Benjamins.
Kim J and Davis C (2003) Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language* **49**, 484–499.
Klein EC (1995) Second vs. third language acquisition: Is there a difference? *Language Learning* **45**, 419–465.
Kwak H-Y (2014) Korean-speaking L2 learners' comprehension of sentences containing numerically quantified NPs and negation in English. *Modern English Education* **15**, 65–84.
Laufer B and McLean S (2016) Loanwords and vocabulary size test scores: A case of different estimates for different L1 learners. *Language Assessment Quarterly* **13**, 202–217.
Lee S (2009) *Interpreting ambiguity in first and second language processing: Universal quantifiers and negation*. Ph.D. dissertation, University of Hawaii, Manoa.
Leeser MJ and Sunderman G (2016) Methodological implications of working memory tasks for L2 processing research. In Granene G, Jackson D and Yilmaz Y (Eds.), *Cognitive individual differences in second language processing and acquisition*, pp. 89–104. Amsterdam: John Benjamins.
Lidz J and Musolino J (2002) Children's command of quantification. *Cognition* **84**, 113–154.
Linck JA, Osthus P, Koeth JT and Bunting MF (2014) Working memory and second language comprehension and production: A meta-analysis. *Psychonomic bulletin & review* **21**, 861–883.
Marsden H (2009) Distributive quantifier scope in English-Japanese and Korean-Japanese interlanguage. *Language Acquisition* **16**, 135–177.
Matuschek H, Kliegl R, Vasishth S, Baayen H and Bates D (2017) Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* **94**, 305–315.
May R (1977) The grammar of quantification. Ph.D. dissertation, MIT university.
May R (1985) *Logical Form*. Cambridge: MIT Press.
Montrul S (2010) Current issues in heritage language acquisition. *Annual Review of Applied Linguistics* **30**, 3–23.
Mueller ST and Piper BJ (2014) The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods* **222**, 250–259.
Musolino J and Lidz J (2003) The Scope of Isomorphism : Turning adults into children. *Language Acquisition* **11**, 277–291.
Musolino J and Lidz J (2006) Why children aren't universally successful with quantification. *Liguistics* **44**, 817–852.
Nespor M, Shukla M and Mehler J (2011) Stress-timed vs. syllable-timed languages. In van Oostendorp M, Ewen CJ, Hume E and Rice K (Eds.), *The Blackwell companion to phonology*, pp. 1147–1159. Oxford: Wiley-Blackwell.
O'Grady W. (2013) The illusion of language acquisition. *Linguistic Approaches to Bilingualism* **3**, 253–285.
O'Grady W, Kwak HY, Lee OS and Lee M (2011) An emergentist perspective on heritage language acquisition. *Studies in Second Language Acquisition* **33**, 223–245.
O'Grady W, Lee M and Kwak H (2009a) Emergentism and Second Language Acquisition. In Ritchie W and Bhatia T (Eds.) *The new handbook of second language acquisition*, pp. 69–88. Bingley: Emerald Press
O'Grady W, Schafer A, Perla J, Lee O-S and Wieting J (2009b) A Psycholinguistic Tool for the Assessment of Language Loss: The HALA Project. *Language Documentation & Conservation* **3**, 100–112.
Paradis M (2008) Language and communication disorders in multilinguals. In Stemmer B and Whitaker HA (Eds.), *Handbook of the neuroscience of language*, pp. 341–349. Amsterdam: Elsevier.
Paradis M. (2009) *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins.
Pesetsky D (1979) *Russian morphology and lexical theory*. Ms., MIT, Cambridge, Mass.
Puig-Mayenco E, González Alonso J and Rothman J (2020) A systematic review of transfer studies in third language acquisition. *Second Language Research* **36**, 31–64.
R Core Team (2014) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
Rothman J (2010) On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian

Portuguese. *IRAL-International Review of Applied Linguistics in Language Teaching* **48**, 245–273.

Rothman J (2011) L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* **27**, 107–127.

Rothman J (2013) Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. In Baauw S, Dirjkoningen F and Pinto M (Eds.), *Romance languages and linguistic theory 2011*, pp. 217–247. Amsterdam: John Benjamins.

Rothman J (2015) Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition* **18**, 179–190.

Rothman J, González Alonso J and Puig-Mayenco E (2019) *Third language acquisition and linguistic transfer*. Cambridge: Cambridge University Press.

Sánchez L and Bardel C (2017) Transfer from an L2 in third language learning: A study on L2 proficiency. In Angelovska T and Hahn A (Eds.), *L3 syntactic transfer: Models, new developments and implications*, pp. 223–250. Amsterdam: John Benjamins.

Schwartz B and Sprouse R (2019) Contra piecemeal transfer: A reply to Slabakova (2017) and Westergaard et al. (2017). Unpublished talk presented at GASLA XV, University of Nevada, Reno, NV, USA.

Slabakova R (2017) The scalpel model of third language acquisition. *International Journal of Bilingualism* **21**, 651–665.

Slabakova R and García Mayo, M Del P. (2015) The L3 syntax–discourse interface. *Bilingualism: Language and Cognition* **18**, 208–226.

Sohn HM (1999) *The Korean language*. Cambridge: Cambridge University Press.

Song HS and Schwartz BD (2009) Testing the fundamental difference hypothesis: L2 adult, L2 child, and L1 child comparisons in the acquisition of Korean Wh-constructions with negative polarity items. *Studies in Second Language Acquisition* **31**, 323–361.

Ullman MT (2001) The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition* **4**, 105–122.

Westergaard M (2019) Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research*, doi: 10.1177/0267658319884116. Published online by SAGE, November 12, 2019.

Westergaard M, Mitrofanova N, Mykhaylyk R and Rodina Y (2017) Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* **21**, 666–682.

Wen ZE (2016) *Working memory and second language learning: Towards an integrated approach*. Bristol: Multilingual matters.

Wu M and Ionin T (2019) L1-Mandarin L2-English speakers' acquisition of English universal quantifier-negation scope. In Brown MM and Dailey B (Eds.), *Proceedings of the 43rd Boston University Conference on Language Development*. pp. 716–729. Somerville, MA: Cascadilla Press.

Zhou P and Crain S (2009) Scope assignment in child language: Evidence from the acquisition of Chinese. *Lingua* **119**, 973–998.