Using A Real-time Word-Naming Technique to Trace Bilingual Children's Vocabulary Development

Hyunwoo Kim[1], Kitaek Kim[2]

[1]*Yonsei University,* [2]*Seoul National University*

hyunwoo2@yonsei.ac.kr[1], *kitaek@snu.ac.kr*[2]

Abstract

This study tested the validity of a computer-based picture-naming tool, the HALA task, to measure the effect of bilingual children's language exposure in their word retrieval skills. We used this task to address an unresolved issue of whether a relative amount of language input is more strongly associated with a relative versus an absolute language outcome. We measured a proportion of L2 Korean input that bilinguals speaking Russian or Chinese as an L1 received from four types of input sources using a language background questionnaire. We also assessed their word retrieving accuracy and speed using the HALA task. Results from correlation analyses show that the proportion of L2 over L1 input was more strongly correlated with the production skills measured in relative than absolute terms. Also, the effect of language input manifested more prominently in word retrieval speed than in word naming accuracy. Our findings suggest that the picture-naming task successfully captured the effect of relative amount of language input in bilingual word production when language outcomes were measured in relative rather than absolute terms. We propose the task as a tool for tracing the development of bilingual children's word production skills.

Keywords: *HALA task, language exposure, immigrant children, lexical retrieval*

Received 12th December 2021; Revised 22nd December 2021; Accepted 24th December 2021

Journal of Cognitive Science 22(4): 541-566 December 2021

1. INTRODUCTION

Research in the field of bilingualism has provided ample evidence of a systematic relationship between an amount of language experience and vocabulary development in bilingual toddlers (e.g., Hoff, Core, Place, Rumiche, Señor & Parra, 2012; Hurtado, Grüter, Marchman & Fernald, 2014; Marchman, Fernald & Hurtado, 2010; Marchman, Martínez, Hurtado, Grüter & Fernald, 2017; Pearson, Fernández, Lewedeg & Oller, 1997) and preschool children (Rojas, Iglesias, Bunta, Goldstein, Goldenberg & Reese, 2016; Sorenson Duncan & Paradis, 2020a, 2020b; Unsworth, Brouwer, de Bree & Verhagen, 2019). Previous studies converge in the findings that children exposed to a target language for an extended period tend to know more words in that language. However, despite the instrumental role of language experience in bilingual vocabulary learning and development, there is little consensus regarding specific methods of estimating bilingual children's language experience and using this variable to assess bilingual language facility. One issue remaining unresolved in previous studies is whether bilinguals' language outcomes should include a target language only or consider both languages of a bilingual (Grüter, Hurtado, Marchman & Fernald, 2014). Some studies measured bilingual language outcomes in a relative term, namely the proportion of language performance in one language over the other, while others employed absolute measures of bilingual outcomes by focusing only on target language performance. For an accurate estimation of bilingual vocabulary growth, it is crucial to establish which of these measures best represents bilinguals' word knowledge. The current study aims to address this issue by investigating the extent to which the language outcomes operationalized in relative versus absolute terms explain variability in the speed and accuracy of bilingual children's lexical retrieval during word production.

Another issue arising when measuring bilingual's language outcomes

concerns how one can efficiently capture bilingual vocabulary knowledge. Previous research has focused on either accuracy measures, such as vocabulary size (indicated by the number of words a bilingual produces or reports to know), or online receptive measures, such as word recognition speed in spoken language comprehension in young children (e.g., Allen, Genesee, Fish & Crago, 2002; De Houwer, 2007; Marchman et al., 2010; Marchman et al., 2017; Simos, Sideridis, Mouzaki, Chatzidaki & Tzevelekou, 2014; Thordardottir, 2019; Zahar, Cobb & Spada, 2001). Given that lexical proficiency includes both abilities to comprehend and produce target words in a rapid and precise manner (Johnson, Paivio & Clark, 1996), studies investigating effects of language exposure on bilingual lexical development need to consider both accuracy and word retrieval speed. In the absence of commonly agreed-upon practices for measuring bilingual word production abilities, this study attempts to provide a practical solution for this problem by adapting a real-time word-naming technique from the Hawai'i Assessment of Language Access (HALA) project (O'Grady, Schafer, Perla, Lee & Wieting, 2009). The HALA task allows for easy and efficient assessment of bilingual vocabulary production abilities by measuring both word naming accuracy and speed via computer software. During the task, participants are presented with a picture of a human body on a computer screen and prompted to name the designated body part as quickly as possible in a target language. In the meantime, the accuracy and the speed of naming the target term are recorded. The underlying assumption of the HALA task is that response times for naming a target object reflect a speaker's ability to access word concepts and retrieve the appropriate words from memory to express the concepts. Furthermore, since human body parts are the basic concepts existing across almost all languages, the task affords comparative analyses of word production ability between the bilinguals' languages (Jo, Kim & Kim, 2021; O'Grady et al., 2009). In this regard, we expect this tool to offer a useful and convenient

means for researchers and educators who intend to trace bilinguals' vocabulary development.

All in all, the current study asks the following questions: (1) Is a relative amount of language input more strongly associated with a language outcome measured in a relative or an absolute term? (2) Does the HALA task offer a useful and convenient means for tracing bilinguals' vocabulary development?

MEASURING BILINGUALS' LANGUAGE INPUT AND OUTPUT

Bilinguals' language exposure profiles are preferentially captured by the relative amount of input a child receives in one language compared to the other. Two types of measuring tools have been particularly useful for quantifying the amount of input – audio-recordings of child-caregiver interactions and parental interviews. Several researchers estimated bilinguals' language input by observing sample recordings from a parent-child interaction for a limited period. For instance, Allen et al. (2002) recorded a 2-hour-long interaction per session between caregivers and their English-Inuktitut bilingual children (aged 1;8 through 3;9) and measured a proportion of input the children received in each language in the sample recordings. They found that the children patterned largely with their parents in the type and amount of code-mixing between English and Inuktitut words, demonstrating a modulating role of language input for bilingual word usage.

Despite its merit of reflecting a bilingual's actual language use, the audio-recording of sample interactions has often been criticized as not fully capturing the dynamic aspects of variation in language input given to bilingual children since it only looks into a small fraction of bilingual language use (Marchman et al., 2017). This problem has recently been resolved in part by the advent of recording technology, which has enabled researchers to make a more precise and naturalistic assessment of bilingual

language input. Marchman et al. (2017) used the new digital recording system called LENA™ (Language ENvironment Analysis), a wearable device that records the child's interactions with caregivers for the whole day. The researchers recorded a minimum of 8 hours of interactions between Spanish-English bilingual children and their caregivers and analyzed the proportion of words from each language in every 5-minute segment. They found that the quantity of language input provided to the children reliably predicted the children's word knowledge and processing speed. While this recording technology opens up a new possibility to obtain large samples of bilingual children's language interactions in a less obtrusive, naturalistic way, it still requires a lot of work until the device becomes fully accessible to most other researchers. Moreover, direct observations may place considerable demands on researchers as they call for substantial time and effort for data collection, transcription, and coding (Grüter et al., 2014).

As an alternative for direct observations, researchers often use interviews and questionnaires that query bilinguals' input in each language (e.g., Paradis, 2011; Pearson et al., 1997; Prevoo et al., 2014). These measures rely on caregivers' reports on the estimated hours of their child interacting in one language relative to the other. For example, Pearson and colleagues (1997) conducted a personal interview to examine a relationship between language exposure and vocabulary learning in English-Spanish bilingual infants aged 8 to 30 months. The researchers obtained a percentage of exposure to each language at home via interviews with parents and found that the reported amount of language input correlated with the number of words that the children knew for each language. Although interview-based methods depend on parents' subjective estimation of language input, findings from studies using these methods have shown close associations between the relative amount of input in one language over the other and the output in bilinguals' language usage.

Despite the different methods of estimating bilingual input, previous

studies have mostly converged on measuring bilingual language input in a relative term (i.e., the proportion of input in one language over the other), rather than in an absolute term (Gathercole & Thomas, 2009; Grüter et al., 2014; Place & Hoff, 2011). This is because a bilingual's language experience is distributed over two languages, and the extent to which a bilingual is exposed to one language compared to the other shapes his or her proficiency in that language (but see De Houwer, 2011, for a different claim). As opposed to the consistent measurement of language input, however, no consensus has emerged regarding how to operationalize bilingual children's language outcomes. Some studies have gauged absolute language performance in a target language (e.g., Hoff et al., 2012), while others have measured relative performance in one language versus the other (e.g., Pearson et al., 1997).

Recently, Grüter et al. (2014) raised the issue of measuring bilinguals' language performance in relative versus absolute terms. Based on their comparison of results from two previous studies (Hurtado et al., 2014; Marchman et al., 2010), the researchers proposed that for the input measured in a relative term, estimates of language exposure could better capture bilinguals' vocabulary size and online processing efficiency when the constructs associated with language outcomes were also measured in a relative term. Specifically, Grüter et al. noted that the study measuring the impact of relative exposure on the relative speed of lexical comprehension (Hurtado et al., 2014) found a more reliable connection between language input and output than the study employing absolute processing efficiency as language outcomes (Marchman et al., 2010). As an explanation for this outcome, Grüter et al. pointed out that absolute measures of language outcomes often fail to capture individual variability in the effects of the relative language input among bilingual children, which diverges substantially depending on their language experience. However, most research has measured language outcomes either in a relative or in an

absolute term. Because of the variability in learner characteristics and experimental methods across those studies, it is difficult to determine which outcome measure better captures the effect of language input in bilinguals' vocabulary proficiency. This problem points to the need for examining the same bilingual population using the same experimental set-up to properly investigate the consequence of employing different measures for language outcomes (i.e., relative versus absolute measures) in bilinguals.

MEASURES OF BILINGUAL'S VOCABULARY KNOWLEDGE

Previous research investigating the relationship between language exposure and vocabulary outcomes in bilingual children have focused either on vocabulary size (e.g., Allen et al., 2002; Marchman et al., 2010) or on processing efficiency in online word comprehension (e.g., Grüter et al., 2014; Marchman et al., 2010; Marchman et al., 2017). For example, Marchman et al. (2010) measured Spanish-English bilingual children's vocabulary size by calculating the number of words that the children were reported to know, which significantly correlated with the number of concepts that the children produced in each language. The researchers also assessed the children's lexical processing ability through the 'looking-while-listening' procedure, which assesses how fast and accurately a child orients eyes toward a target image upon hearing auditory input. Marchman et al. (2010) found that the children tended to initiate an eye-gaze shift more quickly to the target image in response to English words as their relative exposure to English versus Spanish increased.

Despite the well-established link between bilinguals' language experience and their vocabulary size and/or lexical processing ability, less is known about how language experience influences bilinguals' vocabulary growth in production. Lexical proficiency subsumes not only vocabulary size and word comprehension efficiency but also an ability to access lexical representations from memory and retrieve appropriate target words in a

rapid manner (Glaser, 1992; Johnson et al., 1996). In particular, bilinguals' lexical proficiency involves the maintenance of balanced access to words in an interconnected system of the lexicon from multiple languages (De Bot, 2004). Given that managing potential interference from one language while using the other can present considerable processing challenges to bilinguals (Jessner, 2003), efficient lexical access and retrieval for productive use of target words can serve as a good indication of bilinguals' lexical proficiency.

To capture the full scope of the effect of language input on bilinguals' vocabulary growth, the current study uses a real-time word naming technique as a tool for assessing bilingual children's word retrieval speed and accuracy in production. The picture-naming technique has been widely adopted in bilingual research on code-switching and cross-linguistic influence (e.g., Gollan, Montoya, Fennema-Notestine & Morris, 2005; Ramanujan & Weekes, 2020). Yet, this paradigm has hardly been used for investigating the effect of language input on bilingual word production abilities. The picture-naming technique introduced in the current study takes advantage of its full portability and efficiency of measuring a word retrieval ability in a relatively short time (less than 20 minutes), with the help of a computer program along with an audio recorder. This approach can be used with young children, as well as adults, with no literacy skills because the task is implemented entirely in visual presentation.

The real-time word retrieval technique rests on a solid theoretical background of the Weaker Links hypothesis (Gollan, Montoya, Cera & Sandoval, 2008), which proposes that less practice with a second language (L2) can reduce retrieval efficiency of word forms, leading to slower and less accurate lexical retrieval. This hypothesis implies that L2 learners' lexical retrieval may be strengthened with increased exposure to a target language. Based on this theoretical perspective, we investigated how the target language input that bilingual children received from various domains influenced their L2 lexical access and retrieval in production measured

in a relative versus an absolute term. We expect findings from this study to provide unique insight into the effect of bilingual language experience in word production abilities as well as offering a useful tool for tracing bilingual vocabulary development.

2. METHOD

Participants

The bilingual participants were 68 children (41 girls) from immigrant families in South Korea, consisting of 34 Chinese-speaking and 34 Russian-speaking children. Their mean age was 12 (SD = 0.7), and they were enrolled at a local elementary school in Korea at the time of testing. All participants had been born in China or Russian-speaking countries and raised in Chinese- or Russian-dominant home environments until they migrated to Korea with their parents at the mean age of 6.6 (SD = 2.5). Some children born in Uzbekistan, Kazakhstan, and the small provinces of China reported having experience with their regional languages in early childhood. However, they have not been exposed to these languages since they moved to Korea. No children indicated they had received substantial exposure to Korean before their arrival in Korea: Only a few students said they had encountered a small number of Korean words (e.g., names of Korean food) mainly through media in their home country, yet they had little knowledge of Korean before arrival. The mean length of residence in Korea was 2.6 years (SD = 1.2), ranging from 1 to 5 years. Upon entering the school in Korea, they received an intensive Korean language class as well as taking general courses in Korean for the first two years. After completing the intensive course, they were assigned to a regular class where they took the school curriculum with Korean peers.

Participants' Korean proficiency was measured by an oral picture-narration task (Kim & Schwartz, 2020; Song & Schwartz, 2009; elaborated

subsequently). The task scores strongly correlated with their length of stay in Korea (r = .768, p < .001), indicating that longer exposure led to the increased proficiency in Korean. Despite the score variability, the children were considered highly advanced in Korean as the school teacher indicated that the children had little difficulty in listening, speaking, reading, and writing in Korean, and they took all school subjects in Korean.

The two language subgroups – Chinese- and Russian-speaking learners – differed in the mean age, length of stay, and Korean proficiency. As shown in Table 1, the Chinese speakers were significantly younger (t(66) = –2.397, p = .019, Cohen's d = 0.593), had stayed in Korean for a longer period (t(66) = 2.799, p = .007, Cohen's d = 0.679), and had higher proficiency scores (t(66) = 4.425, p < .001, Cohen's d = 1.073) compared to the Russian speakers. Despite these between-group differences, we collapsed them into a single bilingual group in the global analyses since learners' first language (L1) was not counted as a crucial factor in this study. For exploratory purposes, however, we also conducted by-group analyses in addition to the global analyses.

Table 1. Participants' language background

L1	Age (SD)	Years of stay in Korea (SD)	Proficiency scores (SD)
Chinese (n = 34)	11.8 (0.6)	3.1 (1.2)	16.8 (4.7)
Russian (n = 34)	12.2 (0.8)	2.3 (1.1)	11.4 (5.3)

Materials and procedure

A research assistant tested participants individually in a quiet classroom at their school. They completed the following tasks in the same order: (a) language background questionnaire, (b) picture-narration task, and (c) L1/L2 HALA tasks. The overall procedure took approximately 30 minutes.

Materials for the language background questionnaire were adapted from the language history questionnaire used by Li, Sepanski, and Zhao (2006). Survey questions were presented in both L1 and L2. Items consisted of questions asking information of participants' demographic (e.g., age, gender, nationality, etc.) and language background (e.g., native language, length of stay, etc.). We cross-checked the information from participants' responses with the school records. The questionnaire also queried how often the students received input from the ethnic language (L1; Chinese or Russian) and the host language (L2; Korean) in four different domains: home, school, media, and reading. The category home indicated input provided by family members (parents or siblings) at home. Input from school included interactions with teachers and peers at school. Input from media included visual and/or audio signals from TV, radio, and the internet (through a computer or handheld electronic devices). We included this category because language information provided by media is nowadays a primary source of L2 input for bilinguals (Hong, 2010; Pearson, 2007). Finally, reading included input from offline publications, including textbooks, storybooks, magazines, and comics. Participants provided an estimated amount of input they receive per week from each source, separately for their L1 and L2, on a four-point Likert scale from 0 (never use the language) to 3 (use the language quite frequently). Based on this process, each child provided ratings for each of the two languages. For example, when an L1-Russian child used Russian at home most of the time but never spoke or heard Korean, he/she would choose 3 for Russian and 0 for Korean for the home category. If an L1-Chinese child heard Chinese only sparsely from his/her parents at home but used Korean most of the time, he/she would choose 1 for Chinese but 3 for Korean at home. If a child used the L1 with his/her peers at school but not most of the time, then the child would give 2 for the L1 for the school category.

We also implemented the picture-narration task to assess participants'

Korean proficiency. Materials of the task were adopted from those used in Song and Schwartz (2009). Three sets of four pictures were presented on a computer screen, depicting a sequence of daily events (e.g., washing face, eating, and reading a book). On each trial, participants were prompted to describe each picture in chronological order in Korean. Participants' responses were audio-recorded during the task and then transcribed later. We obtained proficiency scores based on complexity and accuracy of each utterance, following the procedure in Song and Schwartz (2009).

Items for the HALA task included 31 photos describing human body parts. The items were divided into two subsets based on their frequency of use, as listed in Table 2.

Table 2. Items for HALA task (translated in English)

Frequency	Item
Higher frequency (k = 17)	back, leg, ear, lips, eye, mouth, face, nose, fingers, shoulder, foot, stomach, hand, teeth, head, tongue, knee
Lower frequency (k = 14)	ankle, forehead, arm, heel, cheek, neck, chin, palm, elbow, thumb, eyebrow, toe, fingernail, wrist

Each bilingual child was tested in both languages, their L1 and L2, with the testing order counterbalanced across participants. Half of the participants first completed the task in their ethnic language (Chinese or Russian) and about three weeks later in Korean, and the order was reversed for the other half. Task items were presented in the frequency order, with the high-frequency items appearing first, followed by the low-frequency items. This ordering served to facilitate participants' responses by reducing cognitive burdens associated with lexical retrieval because naming less frequent before more frequent items can be taxing for L2 learners (O'Grady

et al., 2009). Prior to the main task, participants worked through six practice items, and the research assistant confirmed that they understood the general task procedure.

The task was implemented using Shockwave Flash animation. The flash animation file can play either offline using Adobe Flash Player or online by uploading the file on the flash player platform (http://flashplayer.fullstacks.net/). After providing oral instruction, the research assistant initiated the experiment by clicking the play button on the screen, which triggered a short beep sound, signaling participants to prepare for naming a following image on the screen. The target part of the image was highlighted in a red circle to draw participants' attention to the designated region (see Figure 1). Participants were prompted to name the highlighted region in the respective language as soon as they encountered the picture on the screen. The image remained on the screen for 4000 milliseconds (higher frequency words) or 4500 milliseconds (lower frequency words) after the beep onset until the next trial began. The whole task session was audio-recorded using an iPhone 6S recording application.

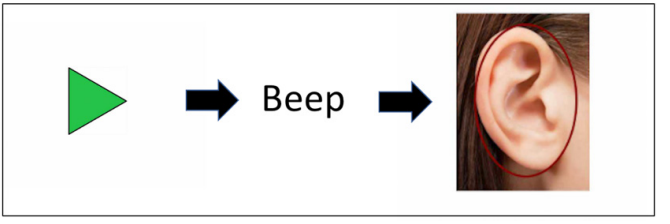

Figure 1. Presentation of an item in the HALA task

Data coding and analyses

For the language input measure, we calculated a relative amount of Korean input for each participant from their responses in the language background questionnaire. Specifically, we divided the ratings of L2 input that the children reported receiving from each type of the four sources (i.e.,

home, school, media, and reading) by the sum of the ratings of L1 and L2 input in each source type. For example, if a child chose '1' for Korean and '3' for Chinese input at school, the relative amount of Korean input that this child received at school was calculated as 0.25 (1/(1+3)).

For the participants' word production skills, we coded their responses in the L1 and L2 HALA tasks in terms of accuracy and word retrieval speed. The accuracy for the L1 task was annotated by trained native speakers of the respective languages, and the accuracy for the L2 task was annotated by an L1 Korean-speaking research assistant and checked by the researchers. Every word accurately responded within the time limit was given one point. Reference to a hypernym for a target word (e.g., a face for the word an eye) was regarded as incorrect, but it was counted as accurate when quickly corrected for the target word (e.g., mentioning a face and then an eye). Accuracy scores in each task were averaged across items. Analogous to the calculation of a relative amount of input, the relative accuracy scores were obtained by dividing the mean accuracy scores from the L2 task by the sum of the mean accuracy scores in the L1 and L2 tasks. The absolute accuracy scores were retrieved from the accuracy means in the L2 task.

Word retrieval speed was calculated as a response time (RT) in milliseconds for every correctly produced word, which was defined as a duration from the beep onset to the onset of the response (O'Grady et al., 2009). RTs were measured using Praat phonetic analysis software (Boersma, 2001). As in the case of the accuracy scores, two types of RTs were obtained: absolute RTs (i.e., RTs in the L2 task) and relative RTs (RTs in the L2 task divided by the sum of RTs in the L1 and L2 tasks).

To address our research question, namely, to what extent the bilingual children's word retrieval ability measured in relative versus absolute terms correlates with the relative amount of L2 input, we conducted Pearson correlation analyses between relative input and relative outcomes and between relative input and absolute outcomes. We then compared

coefficient values obtained from the two comparisons using a Fisher r to z transformation (e.g., Hwang, Jung & Kim, 2020; Kyle & Crossley, 2017) to determine which comparison yields a statistically stronger correlation.

3. RESULTS

Table 3 presents a summary of the data including (a) mean ratings of L2 input and outcomes measured in relative terms, (b) mean ratings for L1 input only and mean scores for L1 outcomes only, and (c) mean ratings for L2 input only and mean scores for L2 outcomes only. Pearson correlation tests revealed significant correlations between the proficiency scores and the input and output measures. The proficiency scores strongly correlated with the word retrieval accuracy (r = .804, p < .001) and with the reaction times (r = –.645, p < .001) in the L2 HALA task, suggesting a close relationship between general L2 proficiency and word production skills. There were also moderate correlations between the proficiency scores and the relative input scores in each of the four domains: home (r = .554, p < .001), school (r = .486, p < .001), media (r = .508, p < .001), and reading (r = .596, p < .001).

Table 3. Statistical descriptions of input and output

	Mean scores for L2 out of L1 and L2 (SD)	**Mean scores for L1 only (SD)**	**Mean scores for L2 only (SD)**
Input type			
Home	0.33 (0.27)	2.39 (0.91)	1.22 (1.04)
School	0.48 (0.22)	2.01 (0.89)	1.92 (0.97)
Media	0.42 (0.31)	2.29 (1.02)	1.66 (1.23)
Reading	0.41 (0.28)	2.06 (1.10)	1.74 (1.10)

Output			
Accuracy	0.43 (0.11)	26.41 (4.44)	20.59 (6.55)
Response time (ms)	0.52 (0.05)	1305.63 (249.12)	1418.13 (267.31)

We scrutinized the relationship between the relative amount of input and the output scores measured in a relative versus absolute term. The input ratings from the four domains were averaged to obtain a composite input score for each participant. Focusing first on accuracy results, the correlation analyses revealed a strong correlation both between the relative amount of input and the accuracy scores measured in a relative term (r = .660, p < .001) and between the relative amount of input and the accuracy scores measured in an absolute term (r = .560, p < .001). Although the correlation coefficient was numerically larger when the accuracy was estimated in a relative than an absolute term (see Figure 1), a Fisher r to z transformation did not show a significant difference between the two coefficient values (z = 0.912; p = .181). These results indicate that the relative amount of Korean input that the bilingual children received was robustly associated with the accuracy for the words that the children produced in the HALA task, no matter whether the accuracy was measured in a relative or an absolute term.

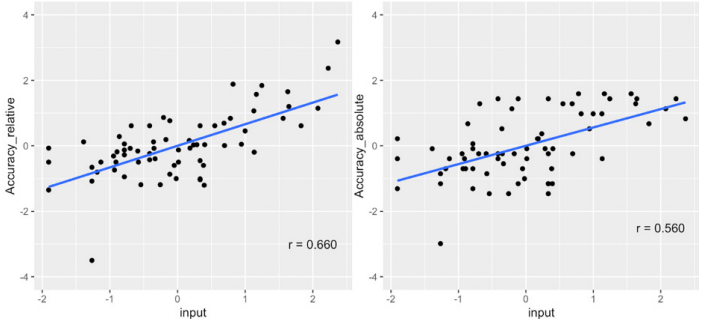

Figure 1. Correlations between input and word retrieval accuracy measured in relative (left) and absolute (right) terms (measures converted to z-scores)

Separate analyses on each L1 group showed comparable results. For the Chinese L1 group, there was a strong correlation between the relative amount of input and the accuracy scores measured in a relative term (r = .689, p < .001) and a moderate correlation between the relative amount of input and the accuracy scores measured in an absolute term (r = .473, p = .005). Similarly, for the Russian L1 group, the relative amount of input correlated moderately with the accuracy scores measured in a relative (r = .454, p = .007) and in an absolute term (r = .454, p = .006).

Turning to the results of RTs, Pearson correlation tests revealed a strong negative correlation between the relative amount of input and the relative RTs (r = –.660, p < .001) and a medium negative correlation between the relative amount of input and the absolute RTs (r = –.371, p = .002), as shown in Figure 2. A Fisher r to z transformation showed a significant difference between the two coefficient values (z = 2.299; p = .011), suggesting that the relative amount of input was more strongly linked to the relative than the absolute RTs.

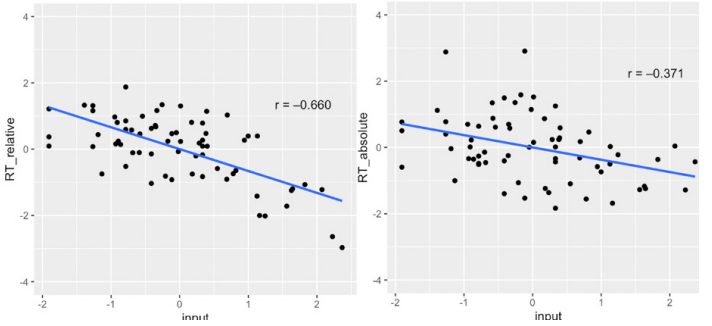

Figure 2. Correlations between input and word retrieval speed measured in relative (left) and absolute (right) terms (measures converted to z-scores)

Separate by-group analyses showed that for the Chinese L1 group, there was a strong negative correlation between the relative amount of input and

the relative RTs (r = –.687, p < .001), but we found no significant correlation between the relative amount of input and the absolute RTs (r = –.241, p = .170). For the Russian L1 group, the relative amount of input correlated moderately with the relative RTs (r = –.415, p = .015), but we found only an approaching significance in the correlation between the relative amount of input and the absolute RTs (r = –.330, p = .057). These results suggest that the relative amount of input was better captured by the language outcomes measured in relative than absolute terms.

4. DISCUSSION

The principal goal of the current study was two-fold: (a) to examine the extent to which language input measured in relative versus absolute terms related to word retrieval accuracy and efficiency in bilingual children, and (b) to confirm the validity of the real-time word naming task to capture the contribution of input in bilingual children's word production skills. To this end, we estimated a relative amount of Korean input from four sources based on the language background questionnaire and measured the bilinguals' word production accuracy and speed in the HALA task.

Results relevant to the comparison between relatively- and absolutely-measured outcomes showed that the relative amount of Korean input was more strongly associated with the production outputs measured in relative than absolute terms, at least in the measurement of reaction times. Although the coefficient of the relative-to-relative comparison was numerically larger than that of the relative-to-absolute comparison for word retrieval accuracy, we did not find any statistical difference. In contrast, we found a statistical difference in the coefficient values when comparing the word retrieval speed between the two ways of comparisons. We thus conclude that the relative amount of input better explained the reaction times measured in relative than absolute terms. These findings are consistent with the claim

that correlations are enhanced when both language input and output in bilinguals are estimated in relative terms (Grüter et al., 2014). As noted by Grüter et al. (2014), the ideal way of probing the input-output link in bilinguals may be measuring both input and output in absolute terms since these measures allow for capturing individual variability in terms of the overall amount of input and word knowledge. However, given the difficulty of obtaining absolute measures of language input, many researchers opted for extrapolating a relative proportion of exposure to one language over the other (e.g., Allen et al., 2002; Gutiérrez-Clellen & Kreiter, 2003; Pearson et al., 1997). In this scenario, Grüter et al. suggested that language outcomes should be consistently measured in relative terms. As they put it, maintaining the measures of bilingual input and output in relative terms helps "understand the role of balance of exposure to two languages in the development of a bilingual lexicon" (p. 24). Our findings supported the validity of this proposal.

In addition to sustaining the Grüter et al.'s claim, our results also provide notable implications in several respects. First, we found the strengthened link between the relatively-measured input and output beyond the domain of word comprehension, which extends to bilinguals' word retrieval ability during production. This implies that relative language exposure plays a crucial role in bilingual word production in addition to vocabulary knowledge and comprehension efficiency. The current results in this regard shed light on the need for measuring language outcomes in relative terms for any future investigations of the input-output relationship in the domain of bilingual word production.

Second, the improved word retrieval ability as a function of increased L2 input suggests that extended exposure to the target language may help bilingual children inhibit potential interference from their L1 word knowledge during L2 word production. It is widely attested that bilinguals experience cross-language conflicts in all representations including the

lexical system (De Bot, 2004). A long line of research has shown that bilinguals activate orthographic, phonological, and semantic information of their L1 word knowledge when using an L2 (e.g., Duyck, Van Assche, Drieghe & Hartsuiker, 2007; Prior, Degani, Awawdy, Yassin & Korem, 2017; Van Assche, Duyck & Brysbaert, 2013). The magnitude of such cross-language activation may diminish as L2 learners have an increased ability to manage interference from a non-target language with cumulative experience with the L2. Aligning with this claim, our results indicate that the children receiving more L2 relative to L1 input were better able to inhibit their L1 lexicon during L2 production, which may have led to more accurate and faster retrieval of the target L2 words in this study.

Third, the strength of the relative-to-relative comparison was more prominent for word retrieval speed than for accuracy. This result lends support for the theoretical motivation of the HALA task, which assumes that language proficiency is associated with word activation levels and access speed. This theoretical approach, formalized in the Weaker Links Hypothesis (Gollan et al., 2008), invokes the frequency of word use as the primary explanation of the facilitated word access and retrieval speed. In other words, more frequent use of language allows for more robust mapping between word concepts and forms in the bilingual mental lexicon, increasing the levels of activation for target words and thus enabling learners to have faster access to lexical information. Similarly, it appears that the bilingual children in our study may have shown faster response times for the target words as a result of more experience with the Korean language with increasing exposure. If word retrieval ability improves with language experience because frequency increases activation levels of words and leads to better retrieval, one would expect more frequent and recent use of target words to have a greater beneficial effect for bilingual children when producing words in an L2.

Fouth, our findings confirm that the HALA task adopted in this study was

appropriate for capturing the effect of language input in bilinguals' word production abilities. The advantages of this task come in several aspects. First of all, the task is inexpensive and easy to use, implemented in the flash animation format, readily available for field workers and teachers without expertise in computer programming and technology. Simply opening the flash file in a personal computer and recording participants' responses, one can easily obtain information of lexical knowledge and processing speed. Furthermore, this task has a potential to be used for an educational purpose—i.e., as a tool for effectively measuring and tracing learners' vocabulary knowledge. This task is cognitively less demanding and less time-consuming compared to other tasks designed for assessing linguistic knowledge or proficiency. Note that other tasks depend largely on learners' literacy skills (e.g., a fill-in-the-blank task by Brown, 1980) or sentence productions (e.g., description of a frog story by Berman & Slobin, 1994, a picture-narration task by Song & Schwartz, 2009). In contrast to these tasks, the HALA task is implemented only in visual presentation, making it possible to administer it to a broad range of participants from young to adult learners at various proficiency levels. From these vantage points, we hope that further studies investigating bilinguals' vocabulary development benefit from utilizing this task.

Finally, we note some limitations of the study. First, we used a small number of items ($k = 31$) with restricted semantic (body parts) and syntactic categories (nouns). Future studies should include a greater number of items in a broader range of semantic and syntactic categories. Second, while the results of this study support Grüter et al.'s (2014) proposal that language outcomes should be consistently measured in relative terms when input is measured in relative terms, an additional study is needed that investigates a relationship between input and output in absolute terms to further our understanding of the input-output link in bilingualism. While obtaining absolute measures of input and output is difficult, using new digital

recording technology, such as wearable audio-recording systems (e.g., Grüter et al., 2014), may enable easier and more accurate measuring of both input and output in absolute terms.

5. CONCLUSION

This study showed that the relative measure of language input in bilingual children from immigrant families was more closely associated with word retrieval accuracy and speed measured in relative than absolute terms. The study also confirmed the contribution of the HALA task in capturing bilingual word production abilities. A helpful direction for further work aimed at elaborating the relation between input and output in bilinguals will be exploring a wider variety of language learning contexts using the technique outlined here. We believe that such cross-study comparisons will help us better understand the roles of different types of language input in guiding bilingual children's word development.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A8028933). Correspondence concerning this article should be addressed to Kitaek Kim (kitaek@snu.ac.kr), the corresponding author.

References

Allen, S. E. M., Genesee, F. H., Fish, S. A., & Crago, M. B. (2002). Patterns of code mixing in English-Inuktitut bilinguals. In M. Andronis, C. Ball, H. Elston & S. Neuvel (Eds.), *Proceedings of the 37th Annual Meeting of the Chicago Linguistic Society* (pp. 171–188). Chicago, IL: Chicago Linguistic Society.

Berman, R. A., & Slobin, D. I. (Eds.). (1994). *Relating events in narrative: A crosslinguistic developmental study*. New York: Psychology Press.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. Glot International, 5, 341–345.

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, 64, 311–317.

De Bot, K. (2004). Introduction: Special issue on language attrition. *International Journal of Bilingualism*, 8, 233–237.

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied psycholinguistics*, 28, 411–424.

De Houwer, A. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review*, 2, 221–240.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 663–679.

Gathercole, V. C. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, 12, 213–237.

Glaser, W. R. (1992). *Picture naming*. Cognition, 42, 61–105.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787–814.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, 33, 1220–1234.

Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, 24, 267–288.

Grüter, T., Hurtado, N., Marchman, V.A., & Fernald, A. (2014). Language exposure and online processing efficiency in bilingual development: relative versus absolute measures. In T. Grüter & J. Paradis (Eds.), *Input and experience in bilingual development* (pp. 15–36). Amsterdam: John Benjamins.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., Parra, M. (2012). Dual

language exposure and early bilingual development. *Journal of Child Language*, 39, 1–27.

Hong, W. P. (2010). Multicultural education in Korea: Its development, remaining issues, and global implications. *Asia Pacific Education Review*, 11, 387–395.

Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. *Bilingualism: Language and Cognition*, 17, 189–202.

Hwang, H., Jung, H., & Kim, H. (2020). Effects of written versus spoken production modalities on syntactic complexity measures in beginning–level child EFL learners. *The Modern Language Journal*, 104, 267–283.

Jessner, U. (2003). A dynamic approach to language attrition in multilingual systems. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 234–246). Clevedon: Multilingual Matters LTD.

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.

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin*, 120, 113–139.

Kim, K., & Schwartz, B. D. (2020). Learnability in the acquisition of the English tough construction by L1-Korean adult and child L2 learners. *Second Language Research. Advance Publication*. https://doi.org/10.1177/0267658320922594

Kyle, K., & Crossley, S. A. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, 34, 513–535.

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, 38, 202–210.

Marchman, V.A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. *Journal of Child Language*, 37, 817–840.

Marchman, V. A., Martínez, L. Z., Hurtado, N., Grüter, T., & Fernald, A. (2017). Caregiver talk to young Spanish−English bilinguals: comparing direct observation and parent−report measures of dual−language exposure. *Developmental Science*, 20, 1–13.

O'Grady, W., Schafer, A. J., Perla, J., Lee, O. S., & Wieting, J. (2009). A

psycholinguistic tool for the assessment of language loss: The HALA project. *Language Documentation & Conservation*, 3, 100–112.

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, 1, 213–237.

Pearson, B. Z. (2007). Social factors in childhood bilingualism in the United States. *Applied psycholinguistics*, 28, 399–410.

Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18, 41–58.

Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Development*, 82, 1834–1849.

Prevoo, M. J., Malda, M., Mesman, J., Emmen, R. A., Yeniad, N., van Ijzendoorn, M. H., & Linting, M. (2014). Predicting ethnic minority children's vocabulary from socioeconomic status, maternal language and home reading input: different pathways for host and ethnic language. J*ournal of Child Language*, 41, 963–984.

Prior, A., Degani, T., Awawdy, S., Yassin, R., & Korem, N. (2017). Is susceptibility to cross- language interference domain specific? *Cognition*, 165, 10–25.

Ramanujan, K., & Weekes, B. S. (2020). Predictors of lexical retrieval in Hindi–English bilingual speakers. Bilingualism: *Language and Cognition*, 23, 265–273.

Rojas, R., Iglesias, A., Bunta, F., Goldstein, B., Goldenberg, C., & Reese, L. (2016). Interlocutor differential effects on the expressive language skills of Spanish-speaking English learners. *International Journal of Speech-Language Pathology*, 18, 166–77.

Simos, P. G., Sideridis, G. D., Mouzaki, A., Chatzidaki, A., & Tzevelekou, M. (2014). Vocabulary growth in second language among immigrant school-aged children in Greece. *Applied Psycholinguistics*, 35, 621–647.

Song, H. S., & Schwartz, B. D. (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.

Sorenson Duncan, T., & Paradis, J. (2020a). How does maternal education influence the linguistic environment supporting bilingual language development in child

second language learners of English? *International Journal of Bilingualism*, 24, 46–61.

Sorenson Duncan, T. & Paradis, J. (2020b). Home language environment and children's second language acquisition: the special status of input from older siblings. *Journal of Child Language*, 47, 982–1005.

Thordardottir, E. (2019). Amount trumps timing in bilingual vocabulary acquisition: Effects of input in simultaneous and sequential school-age bilinguals. *International Journal of Bilingualism*, 23, 236–255.

Unsworth, S., Brouwer, S., de Bree, E., & Verhagen, J. (2019). Predicting bilingual preschoolers' patterns of language development: Degree of non-native input matters. *Applied Psycholinguistics*, 40, 1189–1219.

Van Assche, E., Duyck, W., & Brysbaert, M. (2013). Verb processing by bilinguals in sentence contexts: The effect of cognate status and verb tense. *Studies in Second Language Acquisition*, 35, 237–259.

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review*, 57, 541–572.