Determining the Relationship of Children’s Socioeconomic Background with Vocabulary Development: A Longitudinal Study of Korean Children at 3 and 7 Years of Age

Jeesun Kim, Soo-Jin Kim, Nam Hee Do

*The MARCS Institute of Western Sydney University, Sydney, Australia
bDepartment of Communication Disorders, Korea Nazarene University, Cheonan, Korea
cPSKC Research Team, Korea Institute of Child Care and Education, Seoul, Korea

Correspondence: Soo-Jin Kim, PhD
Department of Communication Disorders, Korea Nazarene University, 48 Wolbong-ro, Seobuk-gu, Cheonan 31172, Korea
Tel: +82-41-570-7978
Fax: +82-41-570-7846
E-mail: sjkim@kornu.ac.kr

Received: April 5, 2018
Revised: May 24, 2018
Accepted: May 24, 2018

This research was supported by the Korea Nazarene University.

Objectives: The current study examined whether early receptive and expressive vocabulary development is related with children’s socioeconomic background by analyzing longitudinal data of Korean children at 3 and 7 years of age. Methods: In general, both vocabulary scores were about 2.5 times higher at 7 years than at 3 years. Socioeconomic background was significantly associated with expressive vocabulary development at age 3, and with the amount of change in the receptive scores between the ages of 3 and 7. While this relationship was maintained in the expressive scores of the 7-year-old, for lower income families the first child tended to develop relatively slowly while the third child developed the quickest. In addition, the results showed that girls’ scores tended to be higher than boys’ scores at 3-year-old, and the difference was maintained in all other scores. Results: There were effects of household income and birth order which interacted. The birth order effect pattern was very similar to the effect of household income. Conclusion: The results are discussed in comparison with previous Korean findings and with international ones.

Keywords: Vocabulary development, Longitudinal study, Korean children, SES, Gender, Birth order

In any given age group, the language ability of individuals varies (Fenson et al., 1994). The issue of why and how such variability occurs, particularly in early childhood, has been at the center of many studies of language development. This is partly because infants’ language ability provides the basic platform upon which subsequent language and academic performance develops (Burchinal, Peisner-Feinberg, Pianta, & Howes, 2002; Hoff, 2013).

Longitudinal studies are pivotal to understanding variability in individual ability, as they provide a basis for examining the factors that modulate developmental progress. Through such studies, evidence can be obtained on how children’s language abilities and competencies develop in relation to individual resources and conditions; the extent to which individual variability emerges early in life; and the trajectories of these abilities in later academic performance and learning. Despite a substantial number of longitudinal, as well as cross-sectional studies, more evidence is needed from a range of cultural, societal and educational settings to identify determining factors.

One of the factors that have been reported to be associated with variability of childhood language development is socioeconomic status (SES), i.e., one’s access to social, financial, and educational resources, and one’s social standing, privileges, and prestige in relation to such access (Duncan, Magnuson, & Votruba-Drzal, 2015; Entwisle & Astone 1994; Mueller &Parcel, 1981). Low SES chil-
dren, in general, perform worse in verbal measures than their peers from high SES families (Arriaga, Fenson, Cronan, & Pethick, 1998; Bowey, 1995; Fernald, Marchman, & Weisleder, 2013; Hart & Risley, 1995, 1999; Rodriguez & Tamis-LeMonda, 2011; Walker, Greenwood, Hart, & Carta, 1994). Childhood SES is typically indexed by parental education, occupation, and/or income (Bradley & Corwyn, 2002; Ensminger & Fothergill, 2003) and it is well-established that maternal education level is an effective predictor of childhood language (and other cognitive) development (Hoff, Laursen, & Bridges, 2012; Magnuson, Sexton, Davis-Kean, & Huston, 2009).

SES-related differences in language abilities emerge from infancy (even in the first year of life; Halle et al., 2009; Vanormelingen & Gillis, 2016) and remain stable or widen over time (Fernald et al., 2013; Halle et al., 2009; Walker et al., 1994). Disparities in SES thus allow predictions to be made about academic trajectories in elementary and secondary school and provide a basis to identify at risk individuals (Burchinal et al., 2002; Entwisle, Karl, Pianta, & Cox, 1999). In this regard, it is worth noting that the most prominent SES-related disparity at an early age is in children’s development of expressive and receptive vocabulary. For example, it has been shown that 18-month-old infants from low SES families have smaller expressive vocabularies compared to their peers from high SES families (Fernald et al., 2013). This difference in expressive vocabulary between low and high SES children is also observed at the age of 3, with low SES children producing half the number of words (Hart & Risley, 1995). These early SES differences have been found across different countries including the United States (Layzer & Price, 2008), the UK (Blanden & Machin, 2010) and Australia (Taylor, Christensen, Lawrence, Mitrou, & Zubrick, 2013), and indicate the important role that word knowledge plays in subsequent language development.

Several theories have been suggested to explain what drives SES-related differences in early language development (Hoff et al., 2012; Weinert, Attig, & Roßbach, 2017). While factors related to an individual’s capacity to learn remain controversial, it is clear that SES-related differences exist in terms of language learning experience and environment. For example, SES is associated with the quantity of language input, the quality of child-caregiver interactions, and differences in language learning opportunities via access to age appropriate materials (Cartmill et al., 2013; Goldin-Meadow et al., 2014). The idea here is that SES related differences in linguistic input translate into differences in language acquisition and development (Hoff-Ginsberg & Shatz, 1982). Moreover, it has been suggested that participating in language exchange through child-directed interaction with caregivers strengthens speech processing skills and, hence, improves the ability to learn from future language encounters (Rowe, 2012; Weisleder & Fernald, 2013). Also, gaining expertise in language allows further autonomous exploration of educational resources available in the environment.

More general SES-related differences in a child’s early environment can also affect language development. Such factors include differential exposure to stressors. For example, it has been suggested that in low SES homes stressors such as violence, household chaos, separation from family members, instability (Evans, 2004) and depression in mothers (Vericker, Macomber, & Golden, 2010) are more prevalent. Children exposed to such stressors are at an increased risk of having reduced language and cognitive skills as well as lower psychological and emotional well-being (Blair & Raver, 2012; Shonkoff et al., 2011). For example, mothers with high levels of stress and depression, talk and interact with their children less often than do low stressed, non-depressed mothers (Lovejoy, Graczyk, O’Hare, & Neuman, 2000). As mentioned above, low levels of language input and poor-quality interactions with children, have the potential to slow vocabulary growth and language development (Pan, Rowe, Singer, & Snow, 2005).

In sum, individual language developments are affected not only by genetic factors that influence learning (Oliver & Plomin, 2007), but also by early experience, as indicated by properties associated with the relation of SES to variation in language outcomes. SES related differences in language development may stem from interactions between learning aptitude and contrasting environments that either drive or delay expressive and receptive language development.

The relationship between children’s language development and parents’ SES has been well established in studies conducted primarily in western countries. Whether such a relationship can be demonstrated in more collectivist, community-focused cultures remains to be confirmed. As such, the current study examined language development in Korean children.

This study used data from a longitudinal study performed by
the Korea Institute of Child Care and Education (KICCE), ‘Panel Study on Korean Children (PSKC).’ PSKC is an ongoing project of KICCE, the aim of which is to establish a nationwide database of Korean children’s growth and development as a function of relevant factors such as environment, family resources and child welfare services and policies. Based on a research model developed in 2006, and a pilot study conducted in 2008, PSKC launched in 2008 and sampled infants (N = 2,150) born in nationwide medical facilities. To reduce sampling bias, a stratified multi-stage sampling method was used so that the participating infants were a fair representation of the young Korean population.

The PSKC was conducted annually until 2015, the year at which the child participants reached elementary school age, and then biennially from 2017 onwards. The data includes measures of children’s linguistic and social development, parents’ psychological characteristics and socioeconomic status, family background, the use of child-care support services, and local community support. In addition to routine data collection, approximately 10% of the participants in the panel study were given an in-depth follow-up at the mean ages of 1, 5, and 7 years; the follow-up included observational study, an interview, and a computer-based test.

The focus of the current study was on the change in receptive and expressive vocabulary size (used as a proxy of language development) of Korean children age 3- to 7-year-old whose SES backgrounds differ. The selection of these ages was made based on studies that report significant SES related differences in development when children are 3 years of age (Brooks-Gunn & Duncan, 1997; Hart & Risley, 1995, 1999). Typically, until age 3, a child’s family is still the primary and often only developmental environment. In contrast, at age 7 social and educational contact occurs daily, exposing children to additional language input and social interaction. Thus, any change across ages 3 to 7 may relate more to the impact of the environment outside of the home. It should be noted that most of the panel study participants (more than 95%) attended kindergarten/preschool during ages 5 to 7.

A special interest of the current study is to determine the extent to which low versus high SES children are at increased risk of delayed language development. This information will be useful for identifying groups at risk, as well as for understanding factors that relate to language development more generally.

**METHODS**

**Participants**

Participants were a nationally representative sample of 1,271 children (52% boys, 48% girls) recruited in the PSKC longitudinal study administered and conducted by KICCE. The mean ages of the children were 3 years (Time_1) and 7 years (Time_2) when the data used in the current study was collected. The infant participants were sampled by using a stratified multi-stage sampling method so that they could be considered an unbiased representation of the young Korean population. For most participants (> 99%), the primary parental caregiver at home was their mother.

**Procedure**

The PSKC data was mainly collected by interviewers trained by KICCE who visited the participants’ households. During household visits, the interviewers conducted surveys, administered developmental tests, and observed and recorded the child participants. A standard version of the Child Behavior Checklist (CBCL; Oh & Kim, 2009) was given to the caregiver to fill out during the interviewers’ visit. Information regarding the child’s temperament and social skills, the mother’s educational background, and household income was collected via a computer assisted personal interview (CAPI).

**Variables of Interest**

Language ability: receptive and expressive vocabulary score

PSKC data on children’s vocabulary proficiency was measured with REVT (Receptive and Expressive Vocabulary Test; Kim, Hong, & Kim, 2009). REVT is a norm-referenced test developed to assess the receptive and expressive vocabulary abilities of Korean speakers ranging from infants to adults; as such, the REVT items consist of vocabulary at different levels of difficulty. The receptive and expressive vocabulary tests consist of 370 items. The receptive subtests (RS) include 185 items that assess comprehension of basic vocabulary. The remaining 185 expressive subsets (ES) consist of a picture labeling task. The total number of correct answers determined the child’s raw score. Children were awarded 1 point for correct items.
Mothers’ educational background

The mothers’ educational background data, which was collected in the 7th year of PSKC, identified 8 levels from ‘no education’ to ‘PhD’. However, considering the sparse distribution at the levels of ‘no education’, ‘elementary and middle school graduate’, and ‘post-graduate degree’, the current study collapsed the data into three levels: (1) high school graduate and lower, (2) college graduate, and (3) university graduate and higher.

Household income

Household income data was collected in the 3rd year of the panel study through a survey where the child’s parents were given an open question and a closed question that required one of three categorical responses. In the present study, household income (Korean won) was classified into (1) upper (monthly house income [MHI] ≥ 4,450,000), (2) middle (2,400,000 < MHI < 4,450,000), or (3) lower (MHI ≤ 2,400,000).

Birth order & gender

Given that language development can be affected by a child’s birth order (Hoff-Ginsberg, 1998) and gender (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Kaushanskaya, Gross, & Buac, 2013), the current study examined the impact of these two variables on RS and ES, alongside household income and mother’s education (Table 1).

RESULTS

Receptive and Expressive Vocabulary Score according to Age

Tables 2-5 present the means and standard deviations of receptive and expressive scores at the age of 3- and 7-year-old according to gender, birth order, household income, and mother’s education levels.

In analyzing the receptive and expressive vocabulary score data at the age of 3- and 7-year-old, we considered that (1) receptive vocabulary is acquired before expressive vocabulary; and (2) vocabulary development at 7 years old is based on vocabulary development at 3 years old. In this regard, in testing whether early vocabulary development varies as a function of household income and mother’s education, we first looked at the receptive scores at 3 years old by conducting a univariate ANOVA. Then, we ran another univariate ANOVA for the expressive scores at the same age with the receptive scores as a covariate. To test vocabulary develop-

Table 1. Number of children that contributed data as a function of household income and maternal education (rows) and gender and birth order (columns)

<table>
<thead>
<tr>
<th>Household income</th>
<th>Mother’s education</th>
<th>Boy</th>
<th>Girl</th>
<th>Grand total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-school</td>
<td>15</td>
<td>12</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>College</td>
<td>12</td>
<td>21</td>
<td>11</td>
<td>44</td>
</tr>
<tr>
<td>University</td>
<td>35</td>
<td>29</td>
<td>9</td>
<td>73</td>
</tr>
<tr>
<td>Total</td>
<td>62</td>
<td>62</td>
<td>25</td>
<td>149</td>
</tr>
<tr>
<td>Middle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-school</td>
<td>57</td>
<td>43</td>
<td>18</td>
<td>118</td>
</tr>
<tr>
<td>College</td>
<td>54</td>
<td>47</td>
<td>16</td>
<td>117</td>
</tr>
<tr>
<td>University</td>
<td>76</td>
<td>72</td>
<td>12</td>
<td>160</td>
</tr>
<tr>
<td>Total</td>
<td>187</td>
<td>162</td>
<td>46</td>
<td>395</td>
</tr>
<tr>
<td>Lower</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-school</td>
<td>24</td>
<td>8</td>
<td>3</td>
<td>35</td>
</tr>
<tr>
<td>College</td>
<td>26</td>
<td>8</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td>University</td>
<td>20</td>
<td>15</td>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>31</td>
<td>10</td>
<td>111</td>
</tr>
<tr>
<td>Grand total</td>
<td>319</td>
<td>255</td>
<td>81</td>
<td>655</td>
</tr>
</tbody>
</table>

Values are presented as mean (SD).
RS = receptive subset; ES = expressive subset.

Table 2. Descriptive statistics of RS/ES by gender and age group

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age of 3 years</th>
<th>Age of 7 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RS</td>
<td>ES</td>
</tr>
<tr>
<td>Boy</td>
<td>32.7 (15.99)</td>
<td>28.8 (13.87)</td>
</tr>
<tr>
<td>Girl</td>
<td>34.4 (15.47)</td>
<td>31.2 (13.36)</td>
</tr>
<tr>
<td>Total</td>
<td>33.5 (15.76)</td>
<td>30.0 (13.68)</td>
</tr>
</tbody>
</table>

Values are presented as mean (SD).
아동의 사회경제적 배경과 어휘발달의 관계 탐색: 한국아동 3세, 7세 종단연구

by the difference in scores for receptive and expressive vocabulary between the ages.

The results for the receptive vocabulary scores at 3 years old showed a significant gender effect ($F_{(1,1217)} = 4.17, p < .05$). Girls had higher scores than boys. The effect of mother education was not significant ($F_{(2,1217)} = 2.01, p = .135$). No other significant effect was found. The results of analyzing the expressive vocabulary scores at 3 years old with the receptive scores as a covariate revealed a significant effect of birth order ($F_{(2,1216)} = 18.01, p < .001$). The scores of the first child tended to be higher than those of the second or third child, and the scores of the second child higher than the third. The

Table 3. Descriptive statistics of RS/ES by birth order and age group

<table>
<thead>
<tr>
<th>Birth order</th>
<th>Age of 3 years</th>
<th>Age of 7 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS</td>
<td>ES</td>
<td>RS</td>
</tr>
<tr>
<td>1st</td>
<td>34.2 (15.22)</td>
<td>32.4 (14.49)</td>
</tr>
<tr>
<td>2nd</td>
<td>33.4 (16.14)</td>
<td>28.7 (12.53)</td>
</tr>
<tr>
<td>3rd &amp; later</td>
<td>31.2 (16.35)</td>
<td>24.6 (12.07)</td>
</tr>
<tr>
<td>Total</td>
<td>33.5 (15.76)</td>
<td>30.0 (13.68)</td>
</tr>
</tbody>
</table>

Values are presented as mean (SD).

RS = receptive subset; ES = expressive subset.

Table 4. Descriptive statistics of RS/ES by house income and age group

<table>
<thead>
<tr>
<th>House income</th>
<th>Age of 3 years</th>
<th>Age of 7 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RS</td>
<td>ES</td>
</tr>
<tr>
<td>Lower</td>
<td>32.9 (16.97)</td>
<td>28.0 (13.88)</td>
</tr>
<tr>
<td>Middle</td>
<td>34.1 (15.81)</td>
<td>30.2 (13.61)</td>
</tr>
<tr>
<td>Upper</td>
<td>32.7 (14.74)</td>
<td>30.8 (13.62)</td>
</tr>
<tr>
<td>Total</td>
<td>33.5 (15.76)</td>
<td>30.0 (13.68)</td>
</tr>
</tbody>
</table>

Values are presented as mean (SD).

RS = receptive subset; ES = expressive subset.
effect of household income was also significant ($F_{(4,1216)} = 3.56, p < .05$). Children from income high families tended to have higher scores. No other significant effect was found.

The analysis of the changes in the receptive scores between 3 and 7 years old showed a birth order effect ($F_{(2,1217)} = 2.99, p < .05$). The first child showed more changes than the second or third child, and the second child showed more changes than the third. The effect of house income was not significant ($F_{(2,1217)} = 2.78, p = .062$). The amount of change in scores tended to positively correlate with the house income level. No other effect was significant. The analysis of the change in the expressive scores between 3 and 7 years old used the change in the receptive scores between 3 and 7 years old as a covariate. This analysis found an interaction between birth order and house income ($F_{(6,1216)} = 2.36, p = .052$). The third born child in the higher income family, followed by the first child in the lower/middle income family, tended to change least whereas the greatest change was observed for the third child born in the lower income family.

Figure 1 provides a visual depiction of the results for the receptive and expressive vocabulary scores as a function of age, gender, birth order and household income. Over all, the effects were small, and few interactions were significant. Tables 2-5 also provide an accessible overview of the effects of these variables on receptive and expressive vocabulary scores (maternal education level is also included).

**DISCUSSION AND CONCLUSION**

The study examined whether the receptive and expressive vocabulary abilities of 3- and 7-year-old children differ as a function of the family SES (in terms of household income and mother’s education) as well as birth order and gender. To summarize the results, the effect of gender was apparent in the receptive scores at age 3 and this gender difference was maintained across all the other scores. The effect of household income and birth order were each apparent in the expressive scores at age 3 and in the amount of change in the receptive scores between the 3 and 7 years old. While these effects were maintained in the development of expressive vocabulary, for lower income families the first child tended to change the least while the third changed the most.

The finding of no effect of mother’s education is surprising, particularly given that it is well-established that maternal education level is a particularly effective predictor of childhood language development (Hoff et al., 2012; Magnuson et al., 2009). The current finding may be related to the rapidly increasing availability of more sophisticated infant-targeted teaching programs at childcare centers and via media (such as television and the Internet). Such programs have the potential to affect children’s language development and the potential for benefit may be greater for children who may have less sophisticated interactions with their mothers who have a lower education level, reducing the relative impact of this factor. In interpreting the current findings, it is also important to note that the lowest educational level included in the current study is basically high-school completion. It could be that the difference between the high-school education and above may be too small to be noticeable in the effect of education on how mothers raise infants (at least in Korea).

The current study showed a significant association between SES (as defined by household income, not by maternal education) and vocabulary development at the ages of 3 and 7. The results are consistent with the findings of other studies (e.g., Hart and Risley, 1995), which reported significant difference in expressive vocabulary between low and high SES children at age of 3; and Fernald et al. (2013), which found that this difference was observed even in 18-month-old infants. Several studies with Korean children have also reported a similar SES effect on early language development (Chang, 2016; Change & Sung, 2011; Lee, Kim, & Kim, 2010; Lee & Kwak, 2008). Further, the current results are consistent with the general claim that SES-related differences in language development emerge from early childhood and remain stable or widen over time (Fernald et al., 2013; Halle et al., 2009; Walker et al., 1994).

With regards to what the observed association of SES with vocabulary development mean, it is not clear the extent to which the association was due to purely different language experiences/environments as a function of SES-related differences. Indeed, one should be particularly cautious here since the current results did not take into account genetic considerations (such as parents’ and children’s IQs). This necessitates further investigation.

The current study found a significant effect of birth order on vocabulary development with the first born more advanced than
the second born, who was more advanced than the later born children. This is consistent with previous findings. For example, Hoff-Ginsberg (1998) observed that firstborn children had more advanced lexical (and grammatical) development than later born children. Kim, Ko, Seo, and Oh (2017) reported a similar effect of birth order in the speech development of Korean children. Hoff-Ginsberg (1998), who observed that later-born children showed more advanced in the social conversation development, suggested that birth order plays a nontrivial role in language development. In this regard, it can be suggested that the early language development of later-born children is ‘different’ rather than ‘disadvantaged’ (Pine, 1995).

In the current study, there was a significant interaction between birth order and family income, i.e., the third-born children in the lower income families showed the most change in the development of expressive vocabulary. It is not clear why this was the case. Here caution needs to be taken in interpreting this difference because the overall number of third-born children was relatively small compared to the first and the second-born. That is, it is hard to tell whether the effect was genuine or due to a smaller sampling size.

Finally, there was a gender effect in vocabulary development at the age of 3 with girl’s scores superior to those of boys. This finding is consistent with several previous studies and supports the general claim that girls’ language abilities develop faster (Chang-Song & Lee, 2006; Kaushanskaya et al., 2013; Leaper, Anderson, & Sanders, 1998; Yoon & Kim, 2004).

In conclusion, the current study of Korean longitudinal data found similarities with studies from other languages and cultures such as English, German, French, etc. That is, Korean children’s vocabulary acquisition at an early age is affected by family SES (as defined by household income) as well as birth order and gender. This study is an important first step in conducting a more advanced investigation of the Korean longitudinal database.

REFERENCES


Entwisle, D. R., & Astone, N. M. (1994). Some practical guidelines for mea-


국문초록

아동의 사회경제적 배경과 어휘발달의 관계 탐색: 한국아동 3세, 7세 종단연구

김지순· 김수진· 도남희

1. The MARCS Institute of Western Sydney University, 2. 나사렛대학교 언어치료학과, 3. 육아정책연구소 아동패널팀

배경 및 목적: 이 연구의 목적은 부모의 사회경제적 지위와 형제순서에 따라 3세와 7세 유아의 수용/표현 어휘발달에 차이가 있는가를 종단적으로 알아보기 위한 것이다. 방법: 아동패널의 3세와 7세 대규모 종단적 발달 자료를 분석하였다. 결과: 모학력에 따른 어휘력의 뚜렷한 차이는 없는 것으로 나타났고 가구의 사회경제적 수준이 높음수록 7세의 점수가 더 높은 것으로 나타났다. 형제순서는 3세와 7세에서 모두 영향을 미쳤다. 논의 및 결론: 이 결과를 국내외 기존 연구들과 비교 논의하였다.

핵심어: 어휘발달, 종단연구, 한국아동, 사회경제적지위(SES), 성별, 형제순서

이 논문은 2018년 나사렛대학교의 연구비 지원에 의해 수행되었음.