Supporting Information

Choi et al. 10.1073/pnas.1706405114

SI Text

Control Analyses. Proportion correct at pretest was above chance for adoptees [$t(28) = 3.9$, $P < 0.01$] and Dutch controls [$t(28) = 3.1$, $P < 0.01$]. Correlations were computed between the adoptees' performance at each test and each of the control factors (i.e., age at time of testing, sex, education, visits to Korea, ratio of length of stay in Korea to length of time since visit to Korea, and number of known languages). Significant negative correlations were found between age at testing and identification accuracy at each of the three tests (pretest, $r = -0.61$, $P < 0.001$; midway test, $r = -0.41$, $P < 0.05$; final test, $r = -0.50$, $P < 0.01$), indicating that, the lower the participant's age, the greater their accuracy on each test. There was also a significant positive correlation between educational level and accuracy at the final test ($r = 0.40$, $P < 0.05$); the higher the participant's level of education, the greater their accuracy on the final test.

To test whether the presence or absence of a relationship with a Korean adoptee may have affected the Dutch controls' performance, ANOVAs on the data from the Dutch control group were conducted with a new variable, relationship (related, unrelated). These showed no significant effects of relationship. The related ($n = 15$) and unrelated control subgroups ($n = 14$) also did not significantly differ in the control factors or in the vocabulary test (Table S5 provides detailed statistics).

Place of Articulation Effects. ANOVAs revealed a significant main effect of place of articulation [$F1(2,112) = 5.4$, $P < 0.01$; $F2(2,216) = 4.0$, $P < 0.05$] and a significant interaction of place of articulation and test [$F1(4,224) = 3.5$, $P < 0.01$; $F2(4,432) = 4.4$, $P < 0.01$]. There was no three-way interaction with group. Analyses following up the two-way interaction showed that, even though there was no significant difference between places of articulation in the pretest, in the later tests, the trained targets received significantly more correct responses than the untrained targets (all P values < 0.05, except for $F2$ comparison of the trained /t/ vs. untrained /p/ at final test, at $P = 0.054$; Table S2 shows proportion correct).

Effects of Sex and of Age at Test. The early-adopted group was more likely to be female and more likely to be younger at test than the later-adopted group. This pattern reflects the trend that adoptive parents select for younger AA and for female adoptees, with this asymmetry increasing with longer adoption record between countries. The same pattern can be observed in the annual Intercountry Adoption reports by the Bureau of Consular Affairs of the US Department of State (https://travel.state.gov/content/adoptionsabroad/en/about-us/statistics.html). Although we found that these two factors conveyed a general advantage in performing our speech sound learning task, we do not believe that explanations internal to our study are to be found. With respect to participant sex, we note that long-term surveys of school performance reveal a stable and persistent female advantage in language subjects (63), that EU statistics (ec.europa.eu/eurostat/statistics-explained/index.php/Tertiary_education_statistics#Gender_distribution_of_participation) reveal that women outnumber men in language course participation across Europe, and that such a "gender gap" is also observed in second-language learners in The Netherlands (64). With respect to participant age across the relatively narrow range of 23–41 y, note that, in a study of a group of more than 3,000 video-game players ranging in age from 16 to 44 y, age-related reduction in performance began at age 24 y and was continuous from then on (65). Although both of these effects were visible in our data, neither was as strong as the effect of being adopted (see *Comparison Analyses: Female Subgroups and Younger Subgroups*).

Comparison Analyses: Female Subgroups and Younger Subgroups. The relative strength of the effects of sex and of age at test in comparison with the effect of being an adoptee were examined in analyses of the female participants only and of the younger participants only. First, the results for female adoptees and female controls were analyzed. ANOVAs on difference scores (midway test minus pretest) were conducted with factor group (adoptees, Dutch control). Results showed that the female adoptees (mean, 0.082) improved more than the female controls (mean, 0.024) between the pretest and midway test, although significance was missed across participants [$F1(1,35) = 3.2$, $P = 0.083$; $F2(1,222) = 15.7$, $P < 0.001$].

Next, we compared the group of younger adoptees against the younger controls in the same way. A median split was used to determine younger participants for each group (adoptees, median age, 31 y, mean age, 28 y, $n = 16$; Dutch controls, median age, 29 y, mean age, 27 y, $n = 15$). ANOVAs showed that the younger adoptees (mean, 0.102) improved significantly more than the younger controls (mean, 0.015) from pretest to midway test [$F1(1,29) = 11.3$, $P < 0.005$; $F2(1,222) = 27.0$, $P < 0.001$].

Comparison of the Early-Adopted Group Against the Whole Control Group. Comparing the early-adopted group against the whole control group in the same way as in the comparison against the selected subgroup of Dutch controls (see main text) led to the same pattern of results [main effect of group, $F1(1,40) = 10.9$, $P < 0.01$; significant group-by-test interaction, $F1(2,80) = 5.4$, $P < 0.01$]. Note that the pretest scores, which were used in the subgroup selection, in fact showed very little variation; the selected control subgroup scored on average 38.4% correct and the remainder of the controls scored 38.2% [$t(27) = 0.088$, $P = 0.93$].

Detailed Procedure for the Childhood-Vocabulary Recognition Test. The test consisted of 10 Korean words that have been shown to be comprehended by 50% of Korean children before the age of 12 mo [ref. 66; using the MacArthur-Bates Communicative Development Inventory (CDI)]. A female native speaker of Korean recorded the words in a clear citation style in a soundproof booth. On each trial, the participants listened to three different recordings of one word. After that, they saw three Dutch words on a computer screen. The participants were asked to indicate which word they thought to be the correct translation (Table S6 lists all materials); they were asked to guess if they did not know. The response options consisted of the 10 correct translations and 20 alternative words. For the selection of the alternative words, English CDI norms were used because no Dutch CDI norms were available at the time of preparation of the present study. The 20 alternative words were Dutch translations of English words that have been shown to be comprehended by 50% of American English children before the age of 12 mo (67). The test was self-paced. The percentage correct was above chance, $t(57) = 7.46$, $P < 0.05$: this might be due to prosody or the onomatopoeic nature of some of the Korean words.

Korean Controls vs. Adoptees and Dutch Controls. To test whether the adoptees and Dutch controls achieved native-like performance in recognizing Korean stops after training, the final-test accuracy of the adoptees and Dutch controls was compared with the accuracy of Korean controls separately for each target consonant. All

comparisons using t tests were significant at $P < 0.001$, showing that the Koreans (with a mean proportion correct across targets of 0.96) significantly outperformed the adoptees and the Dutch controls for all targets.

More Information on Training Procedure. In all sessions, participants were seated in front of a laptop. They heard materials through high-quality headphones and saw instructions and feedback on the screen of the laptop. All responses were given by pressing keys on the laptop keyboard. Presentation software (from the 14 series; Neurobehavioral Systems) was used for constructing and running the experiment.

During the experimenter's four visits, participants received some further exposure to Korean in the form of sentences, short stories, and possible but nonexistent Korean words spoken by several male and female native speakers of Korean, and performed various experimental tasks to be reported elsewhere. This additional exposure was the same for all adoptee and Dutch participants, whereas the native Korean participants did not receive it.

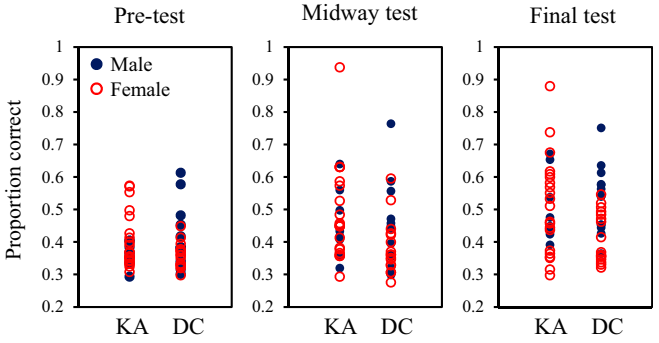

Fig. S1. Individual participants' proportion correct at pretest, midway, and final test as a function of group and sex. DC, Dutch controls; KA, Korean adoptees.

Table S1. Statistics for comparison of Korean adoptees and Dutch control participants on six control variables

		Descriptive statistic			
		Adoptees		Dutch controls	
Variable	Statistical test	*N* (%)	Mean (SD)	*N* (%)	Mean (SD)
Age (y)	$t(56) = -0.24$, $P = 0.13$	29	32 (5)	29	32 (7)
Education*	$\chi^2(3) = 2.21$, $P = 0.53$				
VBO		2 (7)	—	2 (7)	—
MAVO		8 (28)	—	7 (24)	—
HAVO		9 (31)	—	5 (17)	—
VWO		10 (35)	—	15 (52)	—
Visit	$\chi^2(1) = 0.07$, $P = 0.79$				
Yes		16 (55)	—	15 (52)	—
No		13 (45)	—	14 (48)	—
Visit ratio	$t(56) = 0.43$, $P = 0.67$	29	0.01 (0.02)	29	0.01 (0.02)
Language	$t(56) = 0.88$, $P = 0.38$	29	2.8 (1.1)	29	2.6 (1.0)
Sex	$\chi^2(1) = 1.87$, $P = 0.17$				
Female		21 (72)	—	16 (55)	—
Male		8 (28)	—	13 (45)	—

*From lowest to highest level in the Dutch high school system.

Table S2. Mean (SE) correct response proportions for trained alveolar and untrained bilabial and velar targets at pretest, midway, and final test for adoptees and Dutch controls

	Pretest		Midway test		Final test	
Target	Adoptees	Controls	Adoptees	Controls	Adoptees	Controls
Alveolar	0.385 (0.016)	0.383 (0.016)	0.506 (0.025)	0.437 (0.025)	0.534 (0.025)	0.497 (0.025)
Bilabial	0.399 (0.020)	0.396 (0.020)	0.456 (0.026)	0.412 (0.026)	0.495 (0.026)	0.462 (0.026)
Velar	0.391 (0.020)	0.372 (0.020)	0.467 (0.025)	0.384 (0.025)	0.498 (0.025)	0.449 (0.025)

Table S3. Statistics for comparison between early-adopted and later-adopted subgroups on six control variables and on the childhood-vocabulary test

		Descriptive statistic			
		Early-adopted subgroup		Later-adopted subgroup	
Variable	Statistical test	*N* (%)	Mean (SD)	*N* (%)	Mean (SD)
Age (y)	$t(27) = -4.02$, $P < 0.001$	14	28.4 (4.3)	15	34.7 (4.2)
Education*	$\chi^2(3) = 3.37$, $P = 0.34$				
VBO		0	—	2 (13)	—
MAVO		4 (29)	—	4 (27)	—
HAVO		6 (43)	—	3 (20)	—
VWO		4 (29)	—	6 (40)	—
Visit	$\chi^2(1) = 0.91$, $P = 0.34$				
Yes		9 (64)	—	7 (47)	—
No		5 (36)	—	8 (53)	—
Visit ratio	$t(27) = -0.05$, $P = 0.96$	14	0.01 (0.02)	15	0.01 (0.02)
Language	$t(27) = 0.14$, $P = 0.89$	14	2.9 (1.0)	15	2.8 (1.1)
Sex	$\chi^2(1) = 5.66$, $P < 0.05$				
Female		13 (93)	—	8 (53)	—
Male		1 (7)	—	7 (47)	—
Childhood-vocabulary	$t(27) = 1.15$, $P = 0.26$	14	0.50 (0.14)	15	0.43 (0.19)

*From lowest to highest level in the Dutch high school system.

Table S4. Statistics for comparison between early-adopted subgroup and matched Dutch subgroup on six control variables and on the childhood-vocabulary test

		Descriptive statistic			
		Early-adopted subgroup		Matched Dutch subgroup	
Variable	Statistical test	*N* (%)	Mean (SD)	*N* (%)	Mean (SD)
Age (y)	$t(26) = 1.31$, $P = 0.20$	14	28.4 (4.3)	14	31.1 (6.5)
Education*	$\chi^2(3) = 2.89$, $P = 0.41$				
VBO		0	—	1 (7)	—
MAVO		4 (29)	—	2 (14)	—
HAVO		6 (43)	—	4 (29)	—
VWO		4 (29)	—	7 (50)	—
Visit	$\chi^2(1) = 1.29$, $P = 0.26$				
Yes		9 (64)	—	6 (43)	—
No		5 (36)	—	8 (57)	—
Visit ratio	$t(26) = 0.00$, $P = 1.00$	14	0.01 (0.02)	14	0.01 (0.02)
Language	$t(26) = 0.56$, $P = 0.58$	14	2.9 (1.0)	14	2.6 (1.0)
Sex	$\chi^2(1) = 0.00$, $P = 1.00$				
Female		13 (93)	—	13 (93)	—
Male		1 (7)	—	1 (7)	—
Childhood-vocabulary	$t(26) = 0.43$, $P = 0.68$	14	0.50 (0.14)	14	0.48 (0.13)

*From lowest to highest level in the Dutch high school system.

Table S5. Statistics for comparison between related and unrelated control subgroups on six control variables and on the childhood-vocabulary test

		Descriptive statistic			
		Related subgroup		Unrelated subgroup	
Variable	Statistical test	*N* (%)	Mean (SD)	*N* (%)	Mean (SD)
Age (y)	$t(27) = 0.13$, $P = 0.90$	15	32 (6)	14	32 (8)
Education*	$\chi^2(3) = 4.72$, $P = 0.19$				
VBO		1 (7)	—	1 (7)	—
MAVO		5 (33)	—	2 (14)	—
HAVO		4 (27)	—	1 (7)	—
VWO		5 (33)	—	10 (71)	—
Visit	$\chi^2(1) = 0.32$, $P = 0.57$				
Yes		7 (47)	—	8 (57)	—
No		8 (53)	—	6 (43)	—
Visit ratio	$t(27) = -0.37$, $P = 0.71$	15	0.01 (0.02)	14	0.01 (0.02)
Language	$t(27) = -0.29$, $P = 0.78$	15	2.5 (1.1)	14	2.6 (0.9)
Sex	$\chi^2(1) = 0.91$, $P = 0.34$				
Female		7 (47)	—	9 (64)	—
Male		8 (53)	—	5 (36)	—
Childhood-vocabulary	$t(27) = 0.89$, $P = 0.38$	15	5 (1.0)	14	4.6 (1.2)

*From lowest to highest level in the Dutch high school system.

Table S6. Materials used in the word recognition task (and English translations)

	Dutch word		
Korean word	Correct answer	Alternative 1	Alternative 2
mamma (*food*)	eten	bal (*ball*)	neus (*nose*)
kkakkung (*peekaboo*)	kiekeboe	luier (*diaper*)	op (*all gone*)
jjakjjakkung (*clap your hands*)	in je handenklappen	slaaplekker (*good night*)	dansen (*dance*)
manse (*hurray*)	hoera	fles (*bottle*)	mmm lekker (*yum yum*)
mokyok (*bath*)	badje	oh oh (*uh oh*)	baby (*baby*)
swi (*pee*)	plas	sap (*juice*)	kusje (*kiss*)
eungka (*poo*)	poep	melk (*milk*)	schoen (*shoe*)
hajima (*don't do that*)	nietdoen	buiten (*outside*)	beker (*cup*)
jueo (*give*)	geef	hoi (*hi*)	boek (*book*)
jiji (*dirty*)	vies	knuffel (*hug*)	koekje (*cookie*)

Other Supporting Information Files

Dataset S1 (XLS)