Early phonology revealed by international adoptees' birth language retention

Jiyoun Choi[a,b,c,1], Mirjam Broersma[a,d], and Anne Cutler[a,c,1]

[a]Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands; [b]Hanyang Phonetics and Psycholinguistics Lab, Hanyang University, 04763 Seoul, Korea; [c]The MARCS Institute and Australian Research Council Centre of Excellence for the Dynamics of Language, Western Sydney University, NSW 2751, Australia; and [d]Centre for Language Studies, Radboud University, 6500 HD Nijmegen, The Netherlands

Contributed by Anne Cutler, May 19, 2017 (sent for review April 18, 2017; reviewed by Gergely Csibra and Daniel Swingley)

Until at least 6 mo of age, infants show good discrimination for familiar phonetic contrasts (i.e., those heard in the environmental language) and contrasts that are unfamiliar. Adult-like discrimination (significantly worse for nonnative than for native contrasts) appears only later, by 9–10 mo. This has been interpreted as indicating that infants have no knowledge of phonology until vocabulary development begins, after 6 mo of age. Recently, however, word recognition has been observed before age 6 mo, apparently decoupling the vocabulary and phonology acquisition processes. Here we show that phonological acquisition is also in progress before 6 mo of age. The evidence comes from retention of birth-language knowledge in international adoptees. In the largest ever such study, we recruited 29 adult Dutch speakers who had been adopted from Korea when young and had no conscious knowledge of Korean language at all. Half were adopted at age 3–5 mo (before native-specific discrimination develops) and half at 17 mo or older (after word learning has begun). In a short intensive training program, we observe that adoptees (compared with 29 matched controls) more rapidly learn tripartite Korean consonant distinctions without counterparts in their later-acquired Dutch, suggesting that the adoptees retained phonological knowledge about the Korean distinction. The advantage is equivalent for the younger-adopted and the older-adopted groups, and both groups not only acquire the tripartite distinction for the trained consonants but also generalize it to untrained consonants. Although infants younger than 6 mo can still discriminate unfamiliar phonetic distinctions, this finding indicates that native-language phonological knowledge is nonetheless being acquired at that age.

language acquisition | adoption | phonology | language memory

To talk to others, and to understand what others say, we need to know the phonological structure of the language we are using. Phonology embraces, for instance, the single-sound contrasts that distinguish one word from another, or the sequence constraints that apply to sounds; so if we are speaking English, we need to distinguish *rye* from *lie*, and if we hear [*m*] followed immediately by [*b*], we know that they belong to different syllables, as in *somebody*. Neither assumption holds for all languages. Phonological knowledge is highly language-specific.

It is also abstract: not so much knowledge of the sounds and words in the language, but knowledge about those sounds and words, and the rules that govern them. Such knowledge is constantly called into play in adult talking and listening. When talkers must repeat others' utterances, they repeat the sounds that are said, but do so by using their own accent, not by imitating the accent in which the sounds are said (1). If talkers are exposed to novel phoneme sequence constraints in perception, they quickly apply them in production (2). In slips of the tongue, transposed elements accommodate to the new rather than to the originally intended phonetic context, suggesting that abstract rather than phonetically detailed representations were moved (3). Listeners adjust very rapidly to talkers they have never heard before; this adjustment draws on phonological knowledge, in that a talker's unusual pronunciation of a given sound in one phonetic context is quickly generalized to later occurrences of that sound in quite different contexts (4, 5). Likewise, although recognition of previously heard voices is harder in unfamiliar languages than in the native language (6), it is not more difficult in unfamiliar dialects of the native language in which the phonological repertoire is preserved (7). Although listening experience also leaves episodic traces that influence language processing [such as effects of talker familiarity (8) or of frequency (9)], these coexist with everyday use of abstract phonological knowledge.

Children too apply abstract linguistic knowledge. The acquisition of grammar particularly displays this, as young learners overgeneralize rules and, in doing so, produce forms that they have never heard (10). Even infants can apprehend abstract concepts; thus, repetition of structure (e.g., ABB or ABA) is detected in visual sequences by 3-mo-olds (11) and in auditory sequences of trisyllables as early as 2 d of age (12).

However, acquiring phonological structure is a nontrivial task. The brain may be ready to deal with abstract structures, but every aspect of a language's phonology is potentially language-specific, so that it must be learned from the spoken input an infant receives. How much input will yield phonological information, and what prerequisites allow it to be compiled, are questions at the heart of the study of language acquisition in infancy.

For decades, infant speech researchers have assumed that the phonemic repertoire, undeniably central to native language phonology, is largely acquired during the second half of the first year of life. This assumption relies on evidence from speech-discrimination studies in which adult listeners discriminated native-language contrasts but not foreign contrasts, whereas infants of 6–7 mo of age or less discriminated native and foreign contrasts equally well (13, 14). In the second half of the infants' first year, however, an apparent perceptual reorganization occurred, such that, by age 9 mo, discrimination responses had become adult-like [i.e.,

Significance

Dutch adults who, as international adoptees, had heard Korean early in life but had forgotten it learned to identify an unfamiliar three-way Korean consonant distinction significantly faster than controls without such experience. Even adoptees who had been adopted at 3–5 mo of age showed the learning advantage. Thus, early exposure to spoken language, even in the first half-year of life, leaves traces that can facilitate later relearning. Before 6 mo, infants often discriminate foreign-language phonological contrasts better than adults can. This has been widely held to mean that infants younger than 6 mo have no native-language phonological knowledge to capture spoken input. Our findings are significant because they indicate that phonological knowledge is indeed in place before age 6 mo.

Author contributions: J.C., M.B., and A.C. designed research; J.C. performed research; M.B. and A.C. supervised research; J.C. and M.B. analyzed data; and J.C., M.B., and A.C. wrote the paper.

Reviewers: G.C., Central European University; and D.S., University of Pennsylvania.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. Email: jiychoi@hanyang.ac.kr or a.cutler@westernsydney.edu.au.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1706405114/-/DCSupplemental.

significantly better for native than for nonnative contrasts (14, 15)]. This shift from unlimited discrimination to a selective discriminatory benefit for the native contrast inventory occurs rather earlier for vowels than for consonants (16), but the pattern is the same: infants make short work of phoneme acquisition, having the phoneme repertoire essentially in place by age 1 y. They can also by then distinguish between phoneme sequences that are allowable in their language vs. illegal and those that are probable vs. improbable (17, 18).

Also seen in the second half of an infant's first year is evidence of beginning word recognition. Here, infants confront a major challenge in that the words they hear are not presented in isolation, but in normal continuous speech (19); vocabulary acquisition therefore hinges on developing an ability to segment recurring sound patterns from speech. Studies of this ability have suggested that it is absent at 6 mo but in place by approximately 7.5–8 mo (20); individual differences are extensive, and the earlier such segmentation ability appears, the more words are acquired in the following few years (21, 22).

The acquisition of an initial vocabulary and of the native phoneme repertoire in the same general time period made for an attractive scenario in which the two activities were seen as linked, (*i*) in that mastering the phoneme repertoire gave information about which different sequences were actually different words rather than alternative ways of saying the same word (23), (*ii*) in that compiling a repertoire of words gave information about ways in which words could minimally differ (24, 25), or (*iii*) in that the two processes acting together ensured optimal progress toward successful speech processing (26). However, recent studies have pushed the threshold of word recognition to an earlier time point: segmentation of running speech (27, 28) and recognition of words referring to familiar people and concepts (29, 30) have been demonstrated in infants of 6 mo or younger, and neural precursors of word recognition have been observed even at 3 mo (31). These findings cast doubt on any dependence of word recognition on phoneme repertoire possession given that, as has been repeatedly shown, a mature native phoneme inventory bringing reduced sensitivity to foreign contrasts is not in place by 6 mo.

Here we provide evidence of phonological knowledge that is available earlier than suggested by phoneme discrimination studies. The evidence comes from birth language retention by individuals born in one country but adopted at an early age into another country with another language. These international adoptees rapidly acquire the new language they are exposed to (32, 33); after approximately 1 y, they typically command it as well as children who have had only input in that language (34, 35). Their birth language, in contrast, is forgotten. If, at adoption, they could speak in that language, they quickly forget their vocabulary (36–38), and, as adults, they report no language recollection (39, 40).

However, traces of early linguistic experience do remain. Brain responses indicate that birth-language patterns are perceived as linguistically relevant even though the adopted language has no equivalent patterns and native speakers of that language do not show such responses (41). Also, if internationally adopted children are trained on perception of birth-language contrasts with no parallel in their adopted language, they show more rapid initial learning than control learners (42). Adults who were not adopted but were exposed in early childhood to a language spoken by caregivers also show evidence of retention (43), including a rapid learning advantage (44). Adults with a "heritage language"—i.e., a language they have not acquired even though it is spoken by some older family members—experience an advantage when they later enroll in classes in the same language (45, 46). Thus, even though individuals typically report no remaining knowledge of a language they were exposed to early on, varying types of evidence indicate that knowledge has been retained and can influence relearning [note, however, that this advantage, although strong at the outset of training, may lessen with continued instruction (47)].

Drawing on the indications that such influence is particularly detectable in the trajectory of learning upon initial reexposure, we trained adoptee participants (29 Dutch-speaking adults who were adopted from Korea) on birth-language speech sounds (distinct in a way with no parallel in their adopted language of Dutch) and compared their initial learning trajectory to those of closely matched control participants (29 adult Dutch speakers). Our study then focused on the nature of very early phonological knowledge in two ways. First, we assessed whether our participants could generalize the relearned knowledge by testing whether training on one set of speech sounds would lead to improvement in identifying other sounds embodying the same type of distinction. Second, we included as a variable the age at which birth language exposure finished. Based on the aforementioned assumption that a phonemic repertoire would be unknown before 6 mo of age, most prior adoptee studies tested only individuals adopted at age 6 mo or older (e.g., ref. 42), apart from one study of college students of Korean (48), in which seven participants had been adopted from Korea at less than 6 mo of age (some had also had childhood exposure to Korean). We deliberately compared adoptees with enough exposure to have established a phoneme inventory ($n = 15$, adopted at 17 mo or older) against adoptees with less than 6 mo of birth-language exposure ($n = 14$).

Over approximately 1.5 wk, the participants completed an intensive training in identifying three Korean voiceless alveolar stop consonants differing in articulation dynamics, a three-way contrast in which Dutch (like English) has only the one voiceless alveolar stop [*t*]. Before, midway in, and after the training, the participants' performance in identifying the consonants was tested. The test sessions included, besides alveolar stops, voiceless stops embodying the same three-way distinction at two other places of articulation: bilabial ([*p*]) and velar ([*k*]).

Results

Proportions of correct responses were arcsine-transformed for by-subject (*F1*) and by-item (*F2*) repeated-measures ANOVAs, with the between-subject factor group (adoptee or Dutch control), within-subject factors test (pretest, midway, or final), place of articulation (alveolar, bilabial, or velar), and contrast phonation (lenis, fortis, or aspirated). Response times longer than 10 s were excluded (123 trials; 0.27% of data; Dataset S1).

As Fig. 1 shows, adoptees and Dutch controls performed equivalently at pretest, and both groups improved across training. Only the adoptees, however, improved substantially from pretest to midway test; at the midway test, they outperformed the controls. The group difference decreased when training was completed (Fig. S1 shows individual data points). ANOVAs showed a significant effect of test [$F1(2,112) = 35.5$, $P < 0.001$; $F2(2,432) = 137.9$, $P < 0.001$] and an interaction of test with group [$F1(2,112) = 3.2$, $P < 0.05$; $F2(2,432) = 12.4$, $P < 0.001$]. Follow-up pairwise comparisons (Bonferroni-corrected) revealed that the adoptees performed significantly better at the midway test than at the pretest [$F1(1,28) = 14.8$, $P < 0.01$; $F2(1,224) = 71.7$, $P < 0.001$], whereas the Dutch controls' performance did not significantly differ between these tests. Adoptees and controls performed significantly better at the final test than at the midway test (all P values < 0.05). At the midway test, the adoptee/control difference narrowly missed significance [$F1(1,56) = 3.9$, $P = 0.054$; $F2(1,216) = 61.3$, $P < 0.001$]; there was no group difference at the pretest or final test (*SI Text* and Table S1 report further analyses).

Generalization of Learning. Adoptee and control performance at pretest was virtually identical for every place of articulation. The largest improvement with training occurred for the alveolar stimuli used in training (statistics in *SI Text* and Table S2). To test whether generalization of learning of the target contrast to new places of articulation differed across groups, the results for the untrained bilabial and velar targets after training were analyzed separately.

As Fig. 2*A* shows, adoptees performed better than Dutch controls at the midway test, even though significance was again just missed across participants [$F1(1,56) = 3.6$, $P = 0.062$; $F2(1,148) = 36.0$, $P < 0.001$]; separate analyses for each place of articulation showed that the group difference was significant for velar targets [$F1(1,56) = 5.5$, $P < 0.05$; $F2(1,74) = 33.1$,

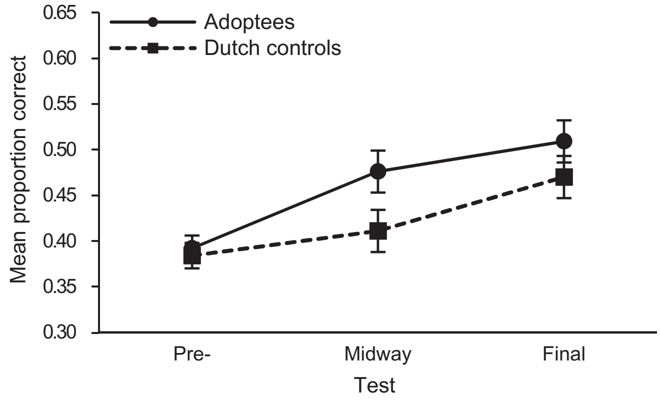

Fig. 1. Mean correct response proportions (across place of articulation and phonation) for adoptees and Dutch controls at pretest, midway, and final test. Error bars show SEs.

$P < 0.001$]. At the final test, the smaller group difference was not statistically significant (Fig. 2*B*).

Adoptive Age. Correlations were computed between adoptive age (AA) and adoptees' performance factors. We first checked, however, whether AA was related to scores in the childhood-vocabulary test (it was not) or to any control factor. Significant positive correlations appeared between AA and age at test ($r = 0.59$, $P < 0.01$) and with sex [$t(27) = 3.38$, $P < 0.01$]: the early-adopted group was younger at test (28 y vs. 35 y) and more likely to be female (13 of 14 vs. 8 of 15; further analyses are provided in Table S3, and further discussion is provided in *SI Text*). With these two factors (age at test, sex) controlled for, all partial correlations between AA and test performance were statistically insignificant.

Fig. 3*A* shows no sign of better performance for the 15 later adoptees who had had longer exposure to Korean compared with the 14 early adoptees; instead, scores were higher for the early-adopted than for the later-adopted group from pretest through the whole training period (note, however, that both groups improved in a similar pattern). ANOVAs with a new factor, AA (early-adopted, later-adopted), and no covariates showed a significant main effect of AA [$F1(1,27) = 7.2$, $P < 0.05$; $F2(1,216) = 166.8$, $P < 0.001$] and no significant interaction involving that factor; when current age and sex were included as covariates in an analysis of covariance, however, this effect of AA disappeared.

Fig. 3*B* compares the early-adopted group to a selected subgroup of the Dutch controls. (This includes the 13 female subjects with the best performances at pretest plus one male subject with the best performance at pretest; this subgroup matched the early-adopted group not only in sex but in all control factors and in vocabulary test results. More comparisons are shown in Table S4.) The early adoptees significantly outperformed this set of controls [$F1(1,26) = 4.3$, $P < 0.05$; $F2(1,216) = 142.3$, $P < 0.001$]. A near-significant group-by-test interaction [$F1(2,52) = 3.1$, $P = 0.055$; $F2(2,432) = 15.0$, $P < 0.001$] reflected no between-group difference at pretest, but better performance by the early-adopted group at midway [mean difference = 0.115, $F1(1,26) = 4.0$, $P = 0.055$; $F2(1,216) = 107.6$, $P < 0.001$] and final tests [mean difference = 0.114, $F1(1,26) = 4.8$, $P < 0.05$; $F2(1,216) = 91.5$, $P < 0.001$]. Comparisons with all controls are detailed in the *SI Text*.

Discussion

Adoptees learned difficult unfamiliar phonetic contrasts on a significantly faster trajectory than control participants could achieve. Stored knowledge, laid down during exposure to their birth language earlier in life, explains how they can have done this. Thus, the results confirmed our initial predictions, motivated by findings in earlier studies: the adoptees had indeed retained phonological knowledge, and effects of this knowledge were revealed primarily in the rapidity with which learning developed.

Surprisingly, however, we found that even a few months of exposure sufficed to establish relevant knowledge that could be drawn upon in relearning. Thus, there is evidence of phonological knowledge that must have been in place before the acquirer was 6 mo old. The comparison we set up, of adults adopted before 6 mo of age vs. adults adopted after 17 mo of age, failed to reveal any knowledge advantage for the latter group. If anything, the earlier-adopted group's learning performance was better, but this was for other reasons, namely that members of the earlier-adopted group were younger at test, and more of them were female; in analyses with these factors as covariates, all differences between the earlier- and later-adopted groups disappeared. Note that, even though being female and being younger were associated with better test performance, being adopted provided by far the greatest advantage: the female adoptees outperformed the female controls, and the younger adoptees outperformed the younger controls (*SI Text*). The important point at issue here is that there was absolutely no sign of the situation that might have been expected on the basis of prior interpretations of the literature, namely that the participants who were adopted earlier would be devoid of phonological knowledge. They were not; they clearly had such knowledge.

The knowledge that our adoptee participants drew upon had, however, not been retained in a form that was consciously accessible or in any way deliberately deployable at the time they were tested. At the beginning of the training regime, the adoptees were unable to do any better than the control participants in identifying the Korean consonants. None of the adoptees showed any recall of Korean childhood vocabulary items, not even the later-adopted group (who should have had

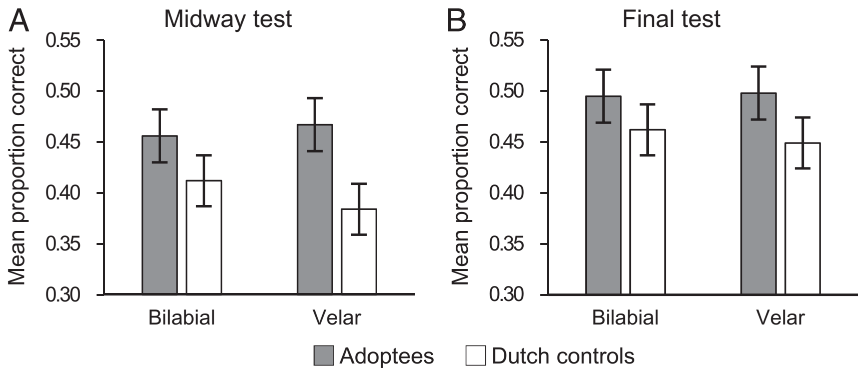

Fig. 2. Mean correct response proportions for untrained bilabial and velar places of articulation for adoptees and Dutch controls at midway (*A*) and final (*B*) tests. Error bars show SEs.

PSYCHOLOGICAL AND COGNITIVE SCIENCES

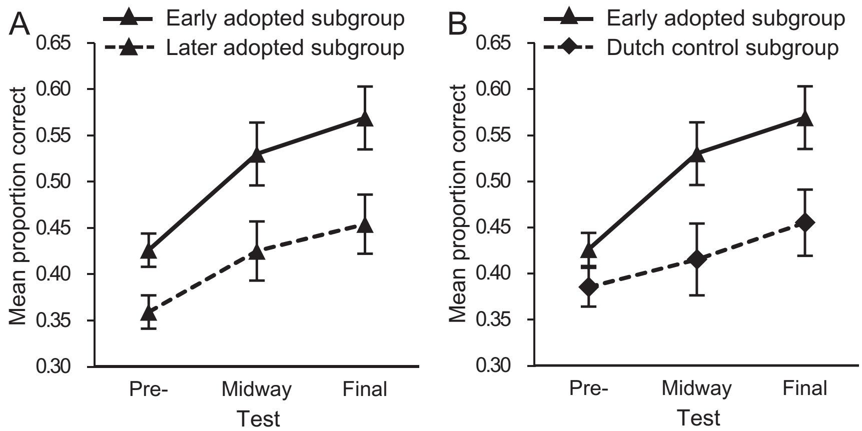

Fig. 3. Mean correct response proportions (across place of articulation and phonation) at pretest, midway, and final test for (*A*) the early-adopted vs. the later-adopted subgroup and (*B*) the early-adopted subgroup vs. a matched Dutch control subgroup. Error bars show SEs.

some ability to talk at the time they were adopted). So what sort of phonological knowledge did they possess?

As described in the Introduction, the evidence from phoneme discrimination experiments strongly suggests that infants have not settled upon the native-language phoneme repertoire before 6 mo of age. There is abundant evidence that, until that age, they successfully discriminate unfamiliar contrasts and do not show the adult indifference to contrasts that are irrelevant to the native language repertoire. In light of that evidence, we therefore reject the notion that the retained knowledge might have been a completed compilation of the Korean phoneme repertoire. The younger group would also not have had a vocabulary of any noteworthy size. The beginnings of successful word segmentation and recognition are seen at approximately 6 mo of age (27–29), but our younger-adopted group left Korea by age 5 mo at the latest; again, the evidence speaks against any serious role for word knowledge at that time.

The question of what kind of knowledge is in place before age 6 mo therefore amounts to one about a type of knowledge that precedes vocabulary initiation (but is not completion of phoneme repertoire compilation) and precedes phoneme repertoire confirmation (but is not word-form learning). One thing we can be reasonably sure of is that the knowledge retained by the adoptees is abstract, i.e., does not consist of accumulated traces of listening experience. The amount of listening experience seems to have played no role in determining whether relevant knowledge was available to our participants, given that our older-adopted group, with more such experience, did not outperform our younger-adopted group. A role for abstract, generally applicable knowledge is also suggested by the fact that the phonetic learning for the one trained place of articulation was generalized to other places of articulation. Further, in specific pursuit of the abstraction question, we also collected production data from the adoptees and the controls (in the form of spoken repetitions of the training stimuli); native speakers identified the adoptees' target utterances more accurately than those of the controls, and also rated the adoptees' productions as more native-like (albeit, on both counts, significantly less acceptable than adult Koreans' utterances) (49). This perception-to-production transfer again suggests the involvement of abstract generalization rather than modality-specific experience.

As already noted, many lines of evidence indicate the capability for abstraction in infant cognition, including sensitivity to repetition of patterns presented visually (11) or auditorily (12); there is also evidence (from 7-mo-old subjects) that acquired abstractions can be further combined in a hierarchical manner (50), a particularly useful achievement for later acquisition of syntactic structure. Abstraction is also involved in children's interpretations of others' gaze as conveying information about the world (51), and in other visual learning performance that cannot be explained by episodic accrual of information (52).

Knowledge about the phonological structure of the environmental language could include its rhythmic patterning; at 2 mo of age, infants discriminate the native language from languages with a different rhythm, but not necessarily from another language with the same rhythm (53). Thus, sensitivity to rhythmic patterning is in place very early. So indeed is sensitivity to other aspects of prosodic structure [infants prefer pauses to be placed between rather than within prosodic phrases (54), for instance], as well as discriminatory preference of licit (e.g., *tap*) over illicit (*tfp*) sequences of consonants and vowels (55, 56). It is thus not unreasonable to propose that the infant notes many systematic properties of speech input before being able to settle on candidate phonemic contrasts and other phonological structures.

Such properties could involve any level of phonological structure. In tone languages, for example, pitch movements are lexically contrastive; infants acquiring such languages show adult-like preference for the native tone inventory over nonnative tone contrasts as early as 4 mo of age (57), suggesting that attention was paid even earlier to the (relatively salient) tonal dimension. Vowel distinctions vary widely across languages, with the modal vowel inventory count standing at five (e.g., in Spanish), but some languages (e.g., English) have many more; this could certainly be a salient property of the input to infants, but whether vowel processing differs in the early months as a function of the size of the environmental vowel inventory is not yet known. Certainly, it is likely that infants are sensitive to the relative patterning of vowels vs. consonants, given that this forms the basis of the rhythmic structure to which they have early sensitivity (53).

In almost every one of the world's phoneme inventories, vowel contrasts are outnumbered by consonant contrasts (58). Consonants differ widely in the type of articulatory gestures they involve [which has consequences for the infant learner (59)], and their acoustic properties largely determine the statistical structure of language-specific speech input (60). Many relevant descriptive generalizations about consonants can be devised and tested against further input. These might encompass questions such as: What articulatory features are deployed? (e.g., Many or few places of articulation? Many or few manners of articulation?); How complex are contrasts? (Are all contrasts binary? Or are there multivalued contrasts?); and Where and how often do consonants occur relative to vowels? (For example, some languages allow only syllables consisting of a consonant plus a vowel.)

Although consonant contrasts have received far more experimental attention in infant speech perception than any other type of contrast, questions about whether they might form classes for the infant learner, or whether their abstract properties might be distinguished, have not been addressed thus far. However, the possible existence of three distinct versions of prevocalic voiceless obstruents, with the distinction expressed in articulatory settings, seems to be a candidate phonological property that infants might note, and store in memory, until such time as they can test it as a useful phonemic construct.

Dutch has only one way of saying prevocalic voiceless obstruents. When required to learn to distinguish three ways of realizing such sounds, our adoptees learned more speedily than the control participants. The knowledge that enabled this might have been as simple as how such an articulation contrast can be realized, or the possibility that contrasts could be multivalued, or both; clearly, however, it is knowledge about possible phonologically relevant properties, and it is laid down quite early on.

This is happening before a vocabulary has been initiated and before the native phoneme inventory has been settled upon. Knowing the phoneme categories can obviously aid compilation of a vocabulary (23), and knowing words can aid definition of the phoneme categories (24, 25). Thus, when the infant brain has matured to the point at which sound/word pairings are being tentatively assessed, for instance, on the basis of statistically likely clusterings of syllables (61), these mutually assisting processes cooperate to produce the rapid progress toward each goal, vocabulary and phoneme repertoire, that we see in the second half of the first year (26). The phonological knowledge that has been built up in the first 6 mo can effectively come into its own in the second 6 mo by delivering useful hypotheses (about constraints on word formation or about potential phoneme repertoire membership) that can be verified or rejected on the basis of the developing vocabulary and inventory. If, as we have concluded, such hypotheses are in abstract form, this may be a crucial basis for the rapidity of the vocabulary and inventory development process; abstraction could make it possible for whole classes of potential phonemic contrasts or potential word structures to be decided upon at once.

That 9–10-mo-old subjects treat foreign-language phonemic contrasts as nonnative and unworthy of discrimination, as adult listeners do, is evidence of the efficiency of the phoneme and vocabulary learning period. That 6–7-mo-old subjects still approach such contrasts with discriminatory interest is evidence that they need to include words in the mix before they can complete the phoneme repertoire learning task. However, their discriminatory ability is not evidence of phonological emptiness, and should never have been viewed in this way. From their earliest days, infants have been amassing invaluable phonological knowledge. As the adoptees described here have shown, this knowledge is firmly laid down and can be of use even if it has been unused for decades.

Methods

Participants. All participants (Korean adoptees, Dutch controls, Korean controls) had at least completed high school. The Korean adoptees were 29 Dutch-speaking adults (21 female, eight male; age range, 23–41 y; mean age, 32 y) recruited through the Dutch Association for Korean Adoptees Arierang and through word of mouth. Their AA by Dutch-speaking families ranged from 3 to 70 mo (5 y 10 mo), with a mean AA of 23 mo (1 y 11 mo). For 14 participants, AA was younger than 6 mo (13 female; range, 3–5 mo; mean, 4 mo; mean age at test, 28 y); for the other 15, AA was older than 17 mo (8 female; range, 1 y 5 mo to 5 y 10 mo; mean, 3 y 3 mo; mean age at test, 35 y). Duration of residence in the Netherlands averaged 30 y 10 mo (range, 23 y 8 mo to 40 y). None had learned Korean after adoption. Thirteen adoptees had never returned to Korea after adoption, whereas the other 16 had made short visits (range, 9–28 d; 12 adoptees visited once, three twice, and one three times).

The 29 Dutch controls (16 female, 13 male; age, 19–47 y; mean age, 32 y) were matched as closely as possible to the adoptees on six potentially relevant factors: (*i*) age at testing, (*ii*) sex, (*iii*) history of visiting Korea (for adoptees, after adoption), (*iv*) mean ratio of length of stay in Korea to time since visit (in days), (*v*) number of languages known, and (*vi*) highest level of completed schooling. Adoptees and controls did not significantly differ in any of these factors (statistics provided in Tables S1 and S5). Half of the Dutch controls (15 of 29) were siblings (nine) or partners (six) of Korean adoptees, and were recruited in a similar way as the adoptees. The other 14 controls had no such relationship. Eight of the latter group, recruited through referrals from family or friends, had made short visits to Korea; the remaining six were recruited from the Max Planck Institute for Psycholinguistics participant pool. No controls had learned Korean.

All participants completed a questionnaire on reasons for participation. Adoptee and control groups answered similarly: "To help research" (29% for adoptees, 32% for controls) or "Out of interest" (34% and 18%, respectively), plus, for controls only, "for partners/siblings" (18%).

A Korean childhood-vocabulary test (detailed in *SI Text* and Table S6) was administered after the main experiments to assess knowledge of Korean. Correct response proportion across adoptees (mean ± SD, 0.46 ± 0.17) and controls (0.48 ± 0.11) did not differ [$t(56) = -0.55$, $P = 0.59$].

Twenty-five native Korean control participants (14 female, 11 male; age range, 27–37 y; mean age, 30 y) were recruited at Hanyang University in Seoul, Korea (control analyses with this group are described in *SI Text*). Neither adoptee nor Dutch control identification performance reached the levels established by this native control group.

No participant reported any hearing loss, uncorrected visual loss, or reading disability. All received a small payment for participating.

Training Stimuli. Twenty-five minimal triplets of Korean disyllabic consonant-vowel-consonant-vowel pseudowords were created. The triplets differed only in word-initial fortis, lenis, or aspirated alveolar stops [t*, t, t^h]. Initial consonant-vowel combinations were the crucial stops plus a vowel: [a], [e], [i], [o], or [u]; second consonant-vowel combinations were [ra], [he], [mi], [tʃo], or [su]. Initial and second consonant-vowel combinations were exhaustively combined, giving 25 triplets in all.

Five male and five female native speakers of standard South Korean (age, 22–33 y) recorded multiple tokens of all 75 items (25 triplets). The items were read one by one in clear citation style in a soundproof booth with a Sennheiser microphone and sampled at 44.1 kHz. The tokens were excised from the recording by using the speech editor PRAAT (62). One token of each item was selected from each talker for training; additional tokens of two triplets from one male talker were used for instructions in all training blocks.

Test Stimuli. The three-way contrast was tested at three places of articulation: alveolar [t*, t, t^h], bilabial [p*, p, p^h], and velar [k*, k, k^h]. For each place of articulation, 25 minimal triplets of Korean consonant-vowel-consonant-vowel pseudowords were included. The triplets for the alveolar stops were taken from the training set; those for the bilabial and velar stops were constructed similarly. A new female native speaker of standard South Korean (age 22) recorded all 225 items (75 items × three places of articulation). One token of each item was selected for the tests. For instructions in all tests, two triplets for each place of articulation were additionally recorded by one of the male training stimuli talkers.

Procedure. Adoptees and Dutch controls completed 13 training blocks and three tests. Training and testing took place in a quiet room at a location chosen by the participants (a home or workplace) over a period of 10–12 d (in 90% of cases, 11 d). In that time, the experimenter (J.C.) visited participants four times (mean interval, 2.3 d). One training block was completed on the first visit, and two blocks were completed on each of the other visits. In each of the three intervisit intervals, participants undertook two further training blocks alone, with laptop and headphones provided by the experimenter. Logs showed that all participants completed the training sessions as instructed. Tests, always conducted in the experimenter's presence, were run before the training on the first visit, after the fourth training block on the second visit, and after the last training block on the fourth visit. Each training block lasted ∼8 min, and each test lasted ∼15 min (more information on the procedure is provided in *SI Text*). Native Korean control participants received no training and performed only a single test.

Training Task. Each of the first 10 training blocks contained stimuli (75 tokens) from a single talker, in each case a different one of the 10 talkers; order of talkers was fixed across participants. The 11th block contained stimuli from the five female talkers (15 tokens per talker), the 12th block from the five male talkers (15 tokens per talker), and the last block contained stimuli from all 10 talkers (7–8 tokens per talker).

The task was three-alternative forced-choice identification. Each training block began with instructions and six practice trials, followed by a break to allow questions. Instructions informed participants that they would hear spoken stimuli, and should listen carefully to the first sound of each stimulus and assign it to one of three categories using the response keys (adjacent on the computer keyboard): "!" (fortis), "@" (lenis), and "#" (aspirated). Example triplets were played twice, with the appropriate symbol on the screen simultaneously highlighted in turn.

Training trials began with a fixation mark in center screen for 400 ms, then a blank screen for 400 ms. One auditory stimulus was then played and participants responded by pressing !, @, or #. Feedback for a correct response was a high-pitched beep and the Dutch word for "good" on the screen in green; for an

PSYCHOLOGICAL AND COGNITIVE SCIENCES

incorrect response, a low-pitched beep and the Dutch for "the correct answer is:" appeared in red, with the correct symbol. There was no response time-out. The 75 stimuli were presented in random order in each training block.

Test Task. In each test, three blocks tested the contrast in alveolar, bilabial, and velar stops, respectively, always in that order. Each block had instructions, six practice trials, an opportunity for questions, and the main test phase, using the three-alternative forced-choice identification procedure as in the training, but without feedback. The same task was used for the Korean controls except that instructions and visual feedback for practice trials were in Korean. This study was carried out in accordance with the recommendations of Radboud University's Ethics Assessment Committee (EAC) Humanities (www.ru.nl/eac-humanities/) and Radboud University's code of academic integrity and conduct (www.ru.nl/english/about-us/our-university/integrity-conduct/) that adhere to European regulations. Informed written consent was obtained from all participants.

ACKNOWLEDGMENTS. The present research was supported by a doctoral fellowship from the Max Planck Society (to J.C.), with additional support from Netherlands Organisation for Scientific Research (NWO)-Veni (M.B.), NWO-Spinoza (A.C.), and, during manuscript preparation, the Australian Research Council Centre of Excellence for the Dynamics of Language (A.C. and J.C.), NWO-Vidi (M.B.), and National Research Foundation of Korea Grant NRF-2016S1A5B5A 01025371 (to J.C.).

1. Mitterer H, Ernestus M (2008) The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition* 109:168–173.
2. Kittredge AK, Dell GS (2016) Learning to speak by listening: Transfer of phonotactics from perception to production. *J Mem Lang* 89:8–22.
3. Fromkin VA (1971) Non-anomalous nature of anomalous utterances. *Language* 47:27–52.
4. McQueen JM, Cutler A, Norris D (2006) Phonological abstraction in the mental lexicon. *Cogn Sci* 30:1113–1126.
5. Cutler A, Eisner F, McQueen JM, Norris D (2010) How abstract phonemic categories are necessary for coping with speaker-related variation. *Laboratory Phonology 10*, eds Fougeron C, Kühnert B, D'Imperio M, Vallée N (de Gruyter, Berlin), pp 91–111.
6. Goggin JP, Thompson CP, Strube G, Simental LR (1991) The role of language familiarity in voice identification. *Mem Cognit* 19:448–458.
7. Johnson EK, Bruggeman L, Cutler A Abstraction and the (misnamed) language familiarity effect. *Cogn Sci*, 10.1111/cogs.12520.
8. Johnson K (1997) Speech perception without speaker normalization: An exemplar model. *Talker Variability in Speech Processing*, eds Johnson K, Mullenix JW (Academic, San Diego), pp 145–165.
9. Pierrehumbert JB (2001) Exemplar dynamics: Word frequency, lenition and contrast. *Frequency Effects and the Emergence of Linguistic Structure*, eds Bybee JL, Hopper P (John Benjamins, Amsterdam), pp 137–157.
10. Marcus GF (1995) Children's overregularization of English plurals: A quantitative analysis. *J Child Lang* 22:447–459.
11. Ferguson B, Waxman SR Visual abstract rule learning by 3- and 4-month-old infants. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, eds Noelle DC, et al. (Cognitive Science Society, Austin, TX), pp 692–697.
12. Gervain J, Macagno F, Cogoi S, Peña M, Mehler J (2008) The neonate brain detects speech structure. *Proc Natl Acad Sci USA* 105:14222–14227.
13. Werker JF, Gilbert JHV, Humphrey K, Tees RC (1981) Developmental aspects of cross-language speech perception. *Child Dev* 52:349–355.
14. Werker JF, Tees RC (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behav Dev* 7:49–63.
15. Werker JF, Lalonde CE (1988) Cross-language speech perception: Initial capabilities and developmental change. *Dev Psychol* 24:672–683.
16. Polka L, Werker JF (1994) Developmental changes in perception of nonnative vowel contrasts. *J Exp Psychol Hum Percept Perform* 20:421–435.
17. Friederici AD, Wessels JMI (1993) Phonotactic knowledge of word boundaries and its use in infant speech perception. *Percept Psychophys* 54:287–295.
18. Jusczyk PW, Luce PA, Charles-Luce J (1994) Infants' sensitivity to phonotactic patterns in the native language. *J Mem Lang* 33:630–645.
19. van de Weijer J (1998) Language input for word discovery. PhD dissertation (Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands).
20. Jusczyk PW, Aslin RN (1995) Infants' detection of the sound patterns of words in fluent speech. *Cognit Psychol* 29:1–23.
21. Junge C, Cutler A (2014) Early word recognition and later language skills. *Brain Sci* 4:532–559.
22. Kooijman V, Junge C, Johnson EK, Hagoort P, Cutler A (2013) Predictive brain signals of linguistic development. *Front Psychol* 4:25.
23. Werker JF, Curtin S (2005) PRIMIR: A developmental framework of infant speech processing. *Lang Learn Dev* 1:197–234.
24. Swingley D (2009) Contributions of infant word learning to language development. *Philos Trans R Soc Lond B Biol Sci* 364:3617–3632.
25. Vihman MM (2017) Learning words and learning sounds: Advances in language development. *Br J Psychol* 108:1–27.
26. Johnson EK (2016) Constructing a proto-lexicon: An integrative view of infant language development. *Annu Rev Linguist* 2:391–412.
27. Johnson EK, Tyler MD (2010) Testing the limits of statistical learning for word segmentation. *Dev Sci* 13:339–345.
28. Johnson EK, Seidl A, Tyler MD (2014) The edge factor in early word segmentation: Utterance-level prosody enables word form extraction by 6-month-olds. *PLoS One* 9: e83546.
29. Tincoff R, Jusczyk PW (1999) Some beginnings of word comprehension in 6-month-olds. *Psychol Sci* 10:172–175.
30. Bergelson E, Swingley D (2012) At 6-9 months, human infants know the meanings of many common nouns. *Proc Natl Acad Sci USA* 109:3253–3258.
31. Friedrich M, Friederici AD (2017) The origins of word learning: Brain responses of 3-month-olds indicate their rapid association of objects and words. *Dev Sci* 20:e12357.
32. Geren J, Snedeker J, Ax L (2005) Starting over: A preliminary study of early lexical and syntactic development in internationally adopted preschoolers. *Semin Speech Lang* 26:44–53.
33. Snedeker J, Geren J, Shafto CL (2007) Starting over: International adoption as a natural experiment in language development. *Psychol Sci* 18:79–87.
34. Clark EA, Hanisee J (1982) Intellectual and adaptive performance of Asian children in adoptive American settings. *Dev Psychol* 18:595–599.
35. Roberts JA, et al. (2005) Language development in preschool-age children adopted from China. *J Speech Lang Hear Res* 48:93–107.
36. Isurin L (2000) Deserted island or a child's first language forgetting. *Biling Lang Cogn* 3:151–166.
37. Glennen S, Masters MG (2002) Typical and atypical language development in infants and toddlers adopted from Eastern Europe. *Am J Speech Lang Pathol* 11:417–433.
38. Nicoladis E, Grabois H (2002) Learning English and losing Chinese: A case study of a child adopted from China. *Int J Biling* 6:441–454.
39. Pallier C, et al. (2003) Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cereb Cortex* 13:155–161.
40. Ventureyra VAG, Pallier C, Yoo H-Y (2004) The loss of first language phonetic perception in adopted Koreans. *J Neurolinguist* 17:79–91.
41. Pierce LJ, Klein D, Chen J-K, Delcenserie A, Genesee F (2014) Mapping the unconscious maintenance of a lost first language. *Proc Natl Acad Sci USA* 111:17314–17319.
42. Singh L, Liederman J, Mierzejewski R, Barnes J (2011) Rapid reacquisition of native phoneme contrasts after disuse: You do not always lose what you do not use. *Dev Sci* 14:949–959.
43. Tees RC, Werker JF (1984) Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Can J Psychol* 38:579–590.
44. Bowers JS, Mattys SL, Gage SH (2009) Preserved implicit knowledge of a forgotten childhood language. *Psychol Sci* 20:1064–1069.
45. Au TK, Knightly LM, Jun S-A, Oh JS (2002) Overhearing a language during childhood. *Psychol Sci* 13:238–243.
46. Oh JS, Jun S-A, Knightly LM, Au TK (2003) Holding on to childhood language memory. *Cognition* 86:B53–B64.
47. Park H-S (2015) Korean adoptees in Sweden: Have they lost their first language completely? *Appl Psycholinguist* 36:773–797.
48. Oh JS, Au TK, Jun S-A (2010) Early childhood language memory in the speech perception of international adoptees. *J Child Lang* 37:1123–1132.
49. Choi J, Cutler A, Broersma M (2017) Early development of abstract language knowledge: Evidence from perception-production transfer of birth-language memory. *R Soc Open Sci* 4:160660.
50. Kovács AM, Endress AD (2014) Hierarchical processing in seven-month-old infants. *Infancy* 19:409–425.
51. Gergely G, Egyed K, Király I (2007) On pedagogy. *Dev Sci* 10:139–146.
52. Csibra G, Gergely G (2009) Natural pedagogy. *Trends Cogn Sci* 13:148–153.
53. Christophe A, Morton J (1998) Is Dutch native English? Linguistic analysis by 2-month-olds. *Dev Sci* 1:215–219.
54. Hirsh-Pasek K, et al. (1987) Clauses are perceptual units for young infants. *Cognition* 26:269–286.
55. Moon C, Fifer WP (1990) Syllables as signals for 2-day-old infants. *Infant Behav Dev* 13: 377–390.
56. Moon C, Bever TG, Fifer WP (1992) Canonical and non-canonical syllable discrimination by two-day-old infants. *J Child Lang* 19:1–17.
57. Yeung HH, Chen KH, Werker JF (2013) When does native language input affect phonetic perception? The precocious case of lexical tone. *J Mem Lang* 68:123–139.
58. Maddieson I (1984) *Pattern of Sound* (Cambridge Univ Press, Cambridge, UK).
59. Best CT, Goldstein LM, Nam H, Tyler MD (2016) Articulating what infants attune to in native speech. *Ecol Psychol* 28:216–261.
60. Guevara Erra R, Gervain J (2016) The efficient coding of speech: Cross-linguistic differences. *PLoS One* 11:e0148861.
61. Swingley D (2005) Statistical clustering and the contents of the infant vocabulary. *Cognit Psychol* 50:86–132.
62. Boersma P, van Heuven V (2001) Speak and unSpeak with PRAAT. *Glot Int* 5:341–347.
63. Cole NS (1997) *The ETS Gender Study: How Females and Males Perform in Educational Settings* (Educational Testing Service, Princeton, NJ).
64. van der Slik FWP, van Hout RWNM, Schepens JJ (2015) The gender gap in second language acquisition: Gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PLoS One* 10:e0142056.
65. Thompson JJ, Blair MR, Henrey AJ (2014) Over the hill at 24: Persistent age-related cognitive-motor decline in reaction times in an ecologically valid video game task begins in early adulthood. *PLoS One* 9:e94215.
66. Pae S, Kwak K (2011) [*Korean MacArthur-Bates Communicative Development Inventories User's Guide and Technical Manual*] (Mind Press, Seoul). Korean.
67. Fenson L, et al. (1994) Variability in early communicative development. *Monogr Soc Res Child Dev* 59:1–173, discussion 174–185.

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.

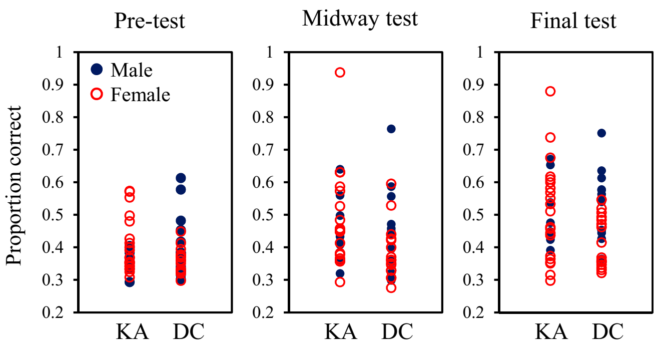

Fig. S1. Individual participants' proportion correct at pretest, midway, and final test as a function of group and sex. DC, Dutch controls; KA, Korean adoptees.

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)