J. Child Lang. **42** (2015), 323–350. © Cambridge University Press 2014
doi:10.1017/S0305000913000627

Can non-interactive language input benefit young second-language learners?*

TERRY KIT-FONG AU
University of Hong Kong

WINNIE WAILAN CHAN
The Hong Kong Institute of Education

LIAO CHENG
Harvard University

LINDA S. SIEGEL
University of British Columbia

AND

RICKY VAN YIP TSO
University of Hong Kong

(*Received* 19 *April* 2012 – *Revised* 28 *March* 2013 – *Accepted* 10 *December* 2013 – *First published online 7 April 2014*)

ABSTRACT

To fully acquire a language, especially its phonology, children need linguistic input from native speakers early on. When interaction with native speakers is not always possible – e.g. for children learning a second language that is not the societal language – audios are commonly used as an affordable substitute. But does such non-interactive input work? Two experiments evaluated the usefulness of audio storybooks in acquiring a more native-like second-language accent. Young children, first- and second-graders in Hong Kong whose native language was Cantonese Chinese, were given take-home listening assignments in a

* We thank the children, parents, and staff of Laichikok Catholic Primary School in Hong Kong for their support of this research. This work was supported by a grant (HKU7400/06H) from the Research Grants Council, Hong Kong SAR (China). We are grateful to a team of dedicated research assistants for data collection, coding, and analysis, to Kathy Shum and Sing-hang Cheung for helping with data analysis, and to Karen Ravn for commenting on earlier drafts. Address for correspondence: Terry Kit-fong Au, Department of Psychology, University of Hong Kong, Centennial Campus, Pokfulam Road, Hong Kong, China. tel: (852) 3917 2383; e-mail: terryau@hku.hk

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

second language, either English or Putonghua Chinese. Accent ratings of the children's story reading revealed measurable benefits of non-interactive input from native speakers. The benefits were far more robust for Putonghua than English. Implications for second-language accent acquisition are discussed.

INTRODUCTION

To fully acquire a language, children need linguistic input from native speakers – and they need to receive it early on (Curtiss, 1977; Johnson & Newport, 1989; Mayberry, 1993). If they do not, the resulting deficits are especially marked in phonology (Flege, Yeni-Komshian & Liu, 1999; Oyama, 1976), leading to noticeable accents that can interfere with communication. Yet because of resource constraints, interaction with native speakers is often unavailable for children learning a second language, especially when the second language is not the societal language (e.g. children learning English in much of Asia and Latin America). Audios are commonly used as an affordable substitute. But does such non-interactive exposure help children acquire a better accent in their second language?

Preverbal infants can learn much about the phonology of the language they hear spoken around them (i.e. their ambient language). Newborns as well as infants aged 0;2 prefer hearing their ambient language to other languages, even when phonetic cues have been removed by low-pass filtering (Dehaene-Lambertz & Houston, 1998; Mehler & Christophe, 1995). They seem able to distinguish their ambient language from other languages on the basis of characteristic intonational and rhythmic patterns. Preverbal infants also can learn much from phonetic cues in their ambient language. They can categorize speech sounds into various consonants and vowels relevant to the language (Eimas, Siqueland, Jusczyk & Vigorito, 1971; Jusczyk, 1997; Werker & Tees, 1984). Such speech sound categories can last for years after exposure to the childhood language has ended (e.g. due to change of caregivers, emigration, or repatriation; Bowers, Mattys & Gage, 2009; Hyltenstam, Bylund, Abrahamsson & Park, 2009; Oh, Au & Jun, 2010; Singh, Liederman, Mierzejewski & Barnes, 2011; Tees & Werker, 1984).

Not all exposure to language is equally useful, though. The nature and amount seem to matter. Consider what happens when infants are exposed to a language that is not their ever-present ambient language. In one foreign-language intervention experiment, American infants aged 0;9 who had a dozen 25-minute sessions of live social interaction with a native speaker of Putonghua (also known as Mandarin) learned to distinguish consonants in that language. But infants who received the same amount of exposure to Putonghua through TV or audios showed little learning

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

(Kuhl, Tsao & Liu, 2003; but see Maye, Werker & Gerken, 2002). These findings suggest that social interaction is necessary in order for infants to map out speech sound categories of a non-ambient language from just a modest amount of input. Perhaps live social interaction facilitates joint attention between adult speakers and infants (Baldwin, 1995; Tomasello & Farrar, 1986) and helps infants recognize the communicative intentions of their partners (e.g. Csibra, 2010).

What about older children? Do they also need social interaction in order to learn any phonology? They do not seem to need such scaffolding to acquire word meanings, which they have been shown to learn from overheard child-directed speech around age 2;0 (e.g. personal pronouns, Oshima-Takane, 1988; novel object and action labels, Akhtar, 2005). On the other hand, children do seem to need social interaction in order to actually learn to speak a language. At least, TV viewing alone is not enough. Hearing-children-of-deaf-parents cannot learn to speak simply by watching TV with the sound on (Sachs, Bard & Johnson, 1981; Todd & Aitchison, 1980). And children cannot learn a foreign language just by watching TV programs in that language (Snow, Arlman-Rupp, Hassing, Jobse, Joosten & Vorster, 1976).

The situation with phonology per se remains unclear though. Can children acquire any phonology from TV viewing or, more generally, from occasional non-interactive input of any kind? The present experiments address this question.

There is some empirical evidence that overhearing live conversations of native speakers during childhood predicts more native-like phonology later. Adults learning Spanish spoke with a more native-like accent if they had overheard Spanish during childhood – even for just several hours a week (mostly in large family parties) – than if they had not. These 'childhood overhearers' spoke English almost exclusively as children, and their spoken Spanish was limited primarily to single words and short phrases (e.g. names for people, Latino food, places, and festivals). Yet in university Spanish language classes, they consistently outperformed classmates with no such childhood experience in several areas: individual phoneme production (e.g. voice onset time of stop consonants), phonological rules (e.g. producing /b, d, g/ as lenited instead of stop consonants in within-word and across-word intervocalic contexts), sentence accent, and story-narrative accent (Au, Knightly, Jun & Oh, 2002; Au, Oh, Knightly, Jun & Romo, 2008; Knightly, Jun, Oh & Au, 2003).

So merely overhearing a language for a few hours a week during childhood predicts more native-like productive phonology of the overheard language. But note that these findings are only correlational in nature, and the causal relation remains unclear. Is the simple fact that children overhear a language really the reason why – then or years later – they can learn to speak it with

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

a better accent? To find out, children's hearing experience has to be manipulated experimentally. Hong Kong offers an invaluable context for such language intervention experiments. Cantonese is the ambient language there for the vast majority of Chinese children, and their exposure to native speakers of other languages – in particular, English and Putonghua – is generally minimal. While these languages are taught in school from kindergarten or first grade on (i.e. by age five or six), the teachers are predominantly non-native speakers due to resource constraints. Outside school, it is likely that children hear native English and Putonghua speakers only randomly (e.g. in brief announcements on trains and buses). TV shows in English or Putonghua are available, but young native speakers of Cantonese-Chinese rarely watch them precisely because of their limited abilities in these non-native languages. It is therefore both feasible and ecologically meaningful to vary experimentally the input children in Hong Kong receive from native speakers of English and Putonghua.

The demonstrated phonological benefits of overhearing live conversations suggest that social interaction may not be necessary for children to make good use of language input (Au *et al*., 2002, 2008; Knightly *et al*., 2003), but does the input have to come from live speakers? From preschool years to early school years, children become more experienced with audio and video media (e.g. TV, videos, music CDs); they also make considerable gains in cognitive abilities (Anderson & Hanson, 2010) and in attention control and self-regulation (e.g. Mezzacappa, 2004; Mischel & Ayduk, 2002). Together, these factors suggest that language input from audio and video media may be enough to help young school-aged children acquire a better second-language accent. The present research therefore focused on that possibility.

In two experiments, we examined potential benefits of non-interactive exposure for young second-language learners through a listening-enrichment program. Two areas of language acquisition were especially relevant in designing this program: (i) the hypothesized sensitive period of language acquisition and (ii) perceptual training in phonology.

Research on age-onset-of-immersion suggests that young children generally outperform older children, who in turn outperform adults in acquiring important aspects of the language being learned – such as phonology (Flege *et al*., 1999; Oyama, 1976) and morphosyntax (Birdsong & Molis, 2001; Johnson & Newport, 1989). Although some late learners can attain near-native mastery of a second language (Birdsong, 1992; White & Genesee, 1996), they are exceptions. The first five or six years of life seem to have a special status in language acquisition. If immersion begins by age five or six, native-like mastery of the language is often achieved. The later the onset, the less native-like the ultimate competence tends to be.

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

What underlies this young-over-old advantage remains debatable. It could be due to younger learners' greater sensitivity to linguistic input, as suggested by the sensitive period hypothesis of language acquisition (Johnson & Newport, 1989; Kuhl, 2010; Werker & Tees, 2005). But it could also be due to younger learners' greater access to good second-language input (Bialystok & Hakuta, 1994, 1999; Flege, Birdsong, Bialystok, Mack, Sung & Tsukada, 2006). For example, immigrant children generally interact with native speakers of a second language actively and frequently at school and at play, whereas older immigrants are more likely to socialize with people from their country of origin. Until this debate is resolved, however, it seems sensible to explore possible benefits of second-language intervention by checking its effects on children still within the hypothesized sensitive period of acquisition. The present experiments therefore recruited children in first and second grades, with ages ranging from five to seven years.

Prior research on perceptual training in phonology also informed our research design. Even infants aged 20 weeks can engage in vocal imitation (Kuhl & Meltzoff, 1982), and infants aged 0;10 babble with speech sounds in their ambient language (de Boysson-Bardies, 1993). This link between sensory and motor experience seems to persist into adulthood. A case in point: many Japanese adults cannot reliably grasp the English /r/–/l/ contrast, perhaps because the Japanese 'r' (e.g. *Hi̲roshi*) sounds like a blend of the English /r/ and /l/. Nonetheless, high-variability perceptual training (with a wide variety of training words recorded by multiple native speakers) improves Japanese adults' production as well as perception of the English /r/ and /l/ (Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999). Analogous benefits of perceptual training have been found in speech therapy for children (e.g. Rvachew, 1994). In the present investigation, then, we emulated the successful high-variability perceptual training in prior research. We invited children to listen to a considerable number of stories, each read aloud by several native speakers in child-directed speech. Hong Kong Cantonese-Chinese speakers in first and second grades were randomly assigned individually to listen to simple stories in either English or Putonghua at home.

Although Cantonese and Putonghua are often described as dialects of Chinese, linguists classify them as distinct languages because native Cantonese speakers do not understand Putonghua on initial contact, and vice versa (Mair, 1991). At the same time, it should be noted that Cantonese and Putonghua are rather similar in their grammars, lexicons, and phonology (Ramsey, 1987). Consider phonology. In both languages, syllables are perceived as taking up roughly the same amount of time (i.e. they are both syllable-timed languages), and there are systematic phonemic mappings, such as /k^h/ and /w/ in Putonghua mapping onto /h/ and /m/ respectively in Cantonese for many words. Moreover, they are both tonal

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

languages, each with several lexical tones. So, while linguistically different languages, Cantonese and Putonghua are perhaps as closely related as Spanish and Portuguese (Chiswick & Miller, 2004).

By contrast, Cantonese and English belong to different language families (Sino-Tibetan versus Indo-European). Unlike the syllable-timed Cantonese and Putonghua, English is a stress-timed language, with the amount of time between consecutive stressed syllables perceived to be on average the same. The linguistic distances between Cantonese and the two target languages may make Putonghua easier than English for Cantonese-speaking children to acquire (e.g. Chiswick & Miller, 2004).

Moreover, Putonghua phonology may be intrinsically easier to acquire than English phonology. For instance, in first language acquisition, Chinese consonants are acquired faster than English ones (So & Dodd, 1995). Moreover, in Putonghua (as well as Cantonese), the syllable structures are relatively simple (V, VC, CV, CVC, where V = vowel, C = consonant), and most words are bisyllabic. By contrast, English allows consonant clusters with as many as three consonants before a vowel (e.g. *strong*) and four after (e.g. *sixths*), and the number of syllables in a word varies considerably, often up to four or five syllables (e.g. *helicopter*, *hippopotamus*). By evaluating comparable enrichment programs in Putonghua and English with native Cantonese-speaking children, we could also explore whether characteristics of the target language would affect children's uptake of non-interactive language input. Our experiments, then, addressed these questions:

1 Can occasional non-interactive exposure to a second language, as in listening to audios of stories read by native speakers, improve young children's accent for that language?
2 Do such benefits, if any, depend on the target language?

EXPERIMENT 1

METHOD

Participants

One hundred and fifty-five children participated with written parental consent (67 boys and 88 girls, first-graders aged from 5;11 to 6;11 at the start of the 18-month experiment). They were recruited from five classrooms in a public elementary school in a middle-class neighborhood in Hong Kong, and the participation rate was 82%. The children spoke Cantonese as their native language and, in accordance with the language education policy in Hong Kong, had begun learning English and Putonghua in first grade. As is the case in most public elementary schools in Hong Kong, their teachers were not native speakers of English or Putonghua. Some children had had

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

prior exposure to these two languages in preschool and kindergarten, but generally not with native speakers. None of the parents were native speakers of either English or Putonghua, according to home-room teachers. The listening enrichment program was offered to all children in these five classrooms as a regular school activity, but the children without parental consent to participate in the research were not assessed in our data collection.

Stimulus materials

(1) Listening-enrichment materials. Altogether, 37 Chinese and 37 English storybooks were selected from two similar series written for beginner readers and published in Hong Kong (*The Magic Box* readers; available at www.greenfieldhk.com). Each story was seven pages long with one sentence/phrase per illustrated page (e.g. *The football game. I can run. I can throw. I can jump. I can kick. I can score the goal. Hooray!* Titles of the books adopted in the enrichment program and a sample of Chinese text are available upon request.). Each Chinese story was audio-recorded by six native speakers of Putonghua from northern Mainland China (3 male, 3 female); each English story was recorded by six native speakers of American-English who had lived both in and outside the US (2 male, 4 female). The readings were all in child-directed speech (Ferguson, 1977; Fernald, 1989). Two or three times a month, every child in this experiment was given a compact disc (CD) containing readings by six native speakers of either Putonghua or English, along with the corresponding storybook in Chinese or English and a listening log. The difficulty level of the stories, as designated in the Chinese and English storybook series, increased gradually over time.

(2) Listening log. To check whether the children actually listened to the audio-recorded stories at home, each child was asked to fill out a 6-item listening-log worksheet while listening to the corresponding take-home CD. On every CD, after each native-speaker's read-aloud of a story, an instruction for one item of the worksheet was given in Cantonese (the children's native language). For example, children were told to draw a square for Item 1 (after listening to the story recorded by the first native speaker), circle the apple among line drawings of three kinds of fruit for Item 2 (after listening to the story recorded by the second native speaker), and so on. Children received a small gift (e.g. an eraser, a pen) for every 30 correct answers in their listening logs. Children's accuracy in filling out the listening logs according to the instructions on the CDs should reflect whether the children had listened to the CDs.

(3) Word-reading assessment at baseline. As noted earlier, the children had some exposure to English and Putonghua prior to this listening-enrichment study. Since assessing the degree of this exposure would have required

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

detailed input from their parents, which would probably have reduced the participation rate, we relied on proxy estimates. For these, we measured children's abilities prior to the enrichment with a read-aloud task in English and Putonghua. The English task was part of the Wide Range Achievement Test (WRAT, Level I); it measured word recognition and was developed for English-speaking children aged five to eleven (Wilkinson, 1993). The Chinese task was part of an assessment of Chinese language abilities for kindergarten and schoolchildren in Hong Kong (HKT-SpLD; Ho, Chan, Lee, Tsang & Luan, 2004), and the children were asked to read aloud in Putonghua. Both tasks used lists of words, ordered by difficulty, that children were asked to read aloud.

Procedure

Each child was randomly assigned individually to either the Putonghua or the English enrichment condition, with the constraint of having a comparable gender ratio in the two conditions in each of the five classrooms. There were seventy-seven children in the Putonghua enrichment condition (44 girls, 33 boys) and seventy-eight children in the English enrichment condition (44 girls, 34 boys). The take-home listening program began in December of the children's first-grade year and ended in June of their second-grade year. There were three modules, mapping roughly onto three consecutive semesters of the school calendar. Take-home listening assignments (each with one story recorded by six native speakers) were distributed by varying schedules because of school activities and holidays, but averaged about two to three stories per month: sixteen stories for Module 1 (from December to June in first grade), nine stories for Module 2 (from September to December in second grade), and twelve stories for Module 3 (from February to June in second grade).

The language enrichment program

The thirty-seven Chinese and thirty-seven English enrichment stories were presented to children as follows:

(1) In-class introduction. The children's regular Putonghua/English teachers introduced each enrichment story in class to make the take-home listening assignments more manageable. A teacher first explained the story in Cantonese, then read it aloud in either Putonghua or English, and finally asked the children to repeat after her line-by-line (i.e. page-by-page) once. To minimize disruption to classroom activities, the children were not separated according to their enrichment condition (Putonghua vs. English) during the in-class story introductions, so such story introductions were not part of the experimental manipulation. The language exposure to native speakers was instead varied by take-home listening assignments.

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

(2) Take-home listening enrichment. On the Friday after the in-class introductions of the Putonghua and English stories, each child received a small bag labeled with the child's name and containing either a Chinese or English storybook (for the Putonghua or English enrichment condition, respectively), a CD with audio-recordings of the story read aloud by six native speakers, and a listening-log worksheet. Children were given the weekend to listen to the story audio (which was about 10 minutes long) and to complete the listening log. The children's home-room teachers collected the listening assignments the following Monday.

This experimental design made it possible to assess benefits specific to the language of enrichment. Because both groups of children received similar treatments in all aspects except the language of enrichment, the two groups served as each other's tailor-made control. If the enrichment program was effective, children who were randomly assigned to the Putonghua enrichment should improve more in their Putonghua accent than in their English accent, and their Putonghua accent improvement should be larger than that of children assigned to the English enrichment. Analogous benefits were predicted for the English enrichment. That is, children who were randomly assigned to the English enrichment should improve more in their English accent than in their Putonghua accent, and their English accent improvement should be larger than that of children assigned to the Putonghua enrichment.

Word-reading assessment at baseline

At the first module pre-test, children were asked to read aloud a list of Chinese words in Putonghua and a list of English words. Children's word readings were audio-recorded with commercially available digital voice recorders (Olympus VN-3100PC) for subsequent coding. Each word-reading task preceded the accent assessment of the respective language.

ACCENT ASSESSMENT

For each module, one of the enrichment stories in each language was chosen for pre-test and post-test to assess improvement in children's Putonghua and English accent for that module. Hence, at pre-test (Time 1), children were asked to read aloud two stories (one each in Putonghua and English). After Module 1 (Time 2) and Module 2 (Time 3), children were asked to read aloud two post-test stories for the just-completed module and two pre-test stories for the next module (again, one each in Putonghua and English). After Module 3, they read only two post-test stories (Time 4). The order of the stories was randomized and counterbalanced across children. For Module 1, the story was used in post-test 19 weeks after it

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

had been a take-home listening assignment. For Modules 2 and 3, the time lag was 6 weeks.

Children were tested individually in a quiet room at school. The pretest for Module 1 was done in the third month of the children's first grade, so their reading abilities in English and Putonghua Chinese were still rather limited. To make the test more feasible, the following testing protocol was used. For each test story, a child was first given 2 or 3 minutes to look at the illustrated storybook. The child was then asked to listen to an audio of the story being read aloud by a native speaker. Finally, the child was asked to read the story aloud, and the reading was audio-recorded with a digital voice recorder for accent rating. The assessment took about 4 minutes per test story. The same testing procedure was used for all four time-points (Time 1 to Time 4) to keep the procedure consistent.

Accent ratings

The children's audio-recordings were edited to yield one sound file per story-reading. There were six test stories (2 languages by 3 modules). Each story was read by 155 children twice (pre-test and post-test), yielding 310 sound files per story. Two native English speakers independently rated all three English test stories. A team of four native Putonghua speakers rated the Putonghua stories, with two independent raters per story. Coders were asked to rate global accents. Such ratings correlated well with acoustical measurement and coding at the phonemic and phonological-rule level in previous research on acquisition of Spanish as a second language (*r*s ranged from ·50 to ·78; Au *et al*., 2002; Knightly *et al*., 2003). Coders were instructed to pay attention to phonological characteristics of English and Putonghua that are challenging for native Cantonese speakers (Chan, 2010; Flege & Wang, 1989; So & Best, 2010): prosody (e.g. intonation contours, stress pattern of multisyllabic words), lexical tones for Putonghua (especially the third tone – the dipping tone), and phonemes (e.g. the English /r/, word-final /t/ and /d/, voicing of /b, d, g/; the Putonghua alveolar /ts, ts^h, s/ and retroflex /tʂ/, /$tʂ^h$/, /ʂ/, /ʐ/ consonants).

For the Putonghua stories, the accent ratings were done on a five-point Likert scale (1 = very strong foreign accent, definitely non-native; 2 = strong foreign accent; 3 = noticeable foreign accent; 4 = slight foreign accent; 5 = no foreign accent, definitely native). This rating scale has been used productively for assessing second-language accents in other languages such as Dutch, Korean, and Spanish (Au *et al*., 2002; Bongaerts, van Summeren, Planken & Schils, 1997; Oh, Jun, Knightly & Au, 2003).

The Putonghua accent scores ranged from 1 to 5 on the 5-point rating scale. By contrast, the English accent scores clustered at the low end, with mostly 1 and 2 and an occasional 3. Consequently, the English accent

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

5-point scale was not very discriminating. To mitigate that problem, half-points were added (i.e. 1, 1·5, 2, 2·5, 3, ...).

For data analysis, an accent score was computed for each story recording by averaging the two accent scores given by the independent raters. For each test story, each accent rater listened to 310 readings produced by all 155 children at both pre-test and post-test. The story recordings were presented in a random order on a computer with the E-Prime software and headphones. Thus the raters did not know whether a specific story-reading had been produced by a child who had heard Putonghua stories or English stories on CDs at home, or whether it had been recorded at pre-test or post-test. Each rater listened to each story recording independently and then entered an accent score into the computer.

RESULTS AND DISCUSSION

Intervention compliance

Children's responses to listening-log worksheets were analyzed to see if they actually listened to the stories in the take-home listening program. Overall, the children listened to the take-home-story CDs very conscientiously. The percentage of correct responses for the thirty-seven worksheets (i.e. 222 items) was 92·9% ($SD = 11·2\%$) in the English enrichment condition and 93·7% ($SD = 5·8\%$) in the Putonghua enrichment condition. The two conditions did not differ significantly ($t(153) = 0·54$, $p > ·5$; all t-tests reported are two-tailed). This pattern of results held in the individual modules as well, so the intervention compliance was generally excellent.

Word-reading abilities at baseline

To estimate the children's prior exposure and home language environment for the two target languages, a response was coded as correct if the word pronunciation was easily recognized by proficient Putonghua and English speakers; the pronunciation did not have to be perfectly native-like. Each response was coded by two independent coders, and the inter-coder agreement was excellent (kappas $> ·95$).

On average, in the Putonghua pre-test, children in the Putonghua and English enrichment conditions correctly read aloud 9·6 words ($SD = 6·2$) and 7·9 words ($SD = 5·8$), respectively. The corresponding averages for the English pre-test were 4·0 words ($SD = 2·2$) and 3·7 words ($SD = 1·6$). The two conditions did not differ significantly in word-reading for either Putonghua ($t(153) = 1·81$, $p > ·07$) or English ($t(153) = 1·00$, $p > ·3$).

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

Inter-rater reliability for accent ratings

The raters of Putonghua accents were four native speakers of Putonghua; they were university students but had no prior formal training in linguistics. The 5-point Likert scale used here was adopted from Bongaerts *et al*.,'s research (1997), which reported reliable global accent ratings given by raters of analogous backgrounds. To achieve inter-rater reliability for each test story in this study, a pair of raters independently coded about 5% of the 310 Putonghua recordings on the 5-point rating scale. Disagreements were discussed to clarify the coding criteria. This process was repeated once or twice until the preliminary Spearman's rho was reasonably high (> ·6). The pair of raters then independently rated all 310 recordings of that test story. The inter-rater reliability between two independent raters for individual test stories in Putonghua was acceptable (Spearman's rhos (309) ranged from ·57 to ·75, ps < ·001).

A similar procedure was used for rating English accents except that (i) half-points were added (i.e. 1, 1·5, 2, 2·5, 3, ...) because most ratings ranged only from 1 to 3, and (ii) all 930 story recordings were rated by the same pair of independent raters, both native speakers of English with university education but no prior formal training in linguistics. The inter-rater reliability for individual test stories between the two independent raters was also acceptable (Spearman's rhos (309) ranged from ·51 to ·77, ps < ·001).

For data analysis, an accent score was computed for each story recording by averaging the two accent scores given by the independent raters. We were mindful that accent ratings were done on ordinal scales and therefore checked whether any disagreements spanned more than two anchor points on the rating scales: 1 = very strong foreign accent, definitely non-native; 2 = strong foreign accent; 3 = noticeable foreign accent; 4 = slight foreign accent; 5 = no foreign accent, definitely native. (Note that the same five anchor points were used for both Putonghua and English accent ratings.) We found that no two raters for any story recording disagreed by more than 1·0 point. Hence, averaging the accent scores for each story recording given by the two independent raters should be fine because the two scores would always fall on either the same anchor point or two adjacent anchor points on the rating scale.

Language enrichment program effects

Table 1 presents the mean pre-test and post-test accent scores for the two enrichment conditions to offer a general idea of the children's Putonghua and English accents. These descriptive statistics, however, need to be interpreted with caution because the Likert scales used for accent ratings were ordinal rather than interval scales. To evaluate the enrichment effects

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

TABLE 1. *Children's mean pre-test and post-test scores (standard deviations) of Putonghua and English accent in Experiment 1*

Putonghua enrichment condition

	Story Accent ($N=77$)	
	Putonghua (With enrichment)	English (Without enrichment)
Pre-test	3·16 (0·70)	1·67 (0·42)
Post-test	3·61 (0·59)	1·91 (0·45)

English enrichment condition

	Story Accent ($N=78$)	
	English (With enrichment)	Putonghua (Without enrichment)
Pre-test	1·62 (0·46)	3·03 (0·66)
Post-test	1·93 (0·48)	3·32 (0·62)

more rigorously, the accent scores were therefore analyzed with ordinal logistic regression models.

Multilevel ordinal logistic regression models were used to fit the data. Post-test accent rating for each of the two target languages, Putonghua and English, was set as the dependent variable in separate analyses. A two-level random intercept structure was conceptualized for the regression procedures, comprising individual accent rating scores (level 1) nested within participants (level 2). In each regression equation, pre-test and post-test accent ratings were specified as ordinal variables, which consisted of five response categories on the Likert scale for Putonghua (i.e. 1, 2, 3, 4, 5) and seven categories for English due to half-points included in the rating scale (i.e. 1, 1·5, 2, 2,5, 3, 3·5, 4; no child scored above 4).

Three models were constructed in each analysis. The baseline model (Model 1)[1] examined the fixed effects of the control variables, which included gender, homework compliance, and pre-test accent rating. In Model 2,[2] two independent categorical variables, namely enrichment condition (Putonghua vs. English) and module (1, 2, and 3) were added, along with the control variables in Model 1. Interaction between enrichment

[1] The function call for Model 1 was study1.chn.fm0 <- clmm(posttest ~ gender + hw_compliance_percent + pretest + (1|subject_no), data = na.omit(study1.chn))

[2] The function call for Model 2 was study1.chn.fm1 <- clmm(posttest ~ gender + hw_compliance_percent + pretest + group_story + module + (1|subject_no), data = na.omit (study1.chn))

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

TABLE 2. *Model estimates for the multilevel ordinal logistic regression analyses on Putonghua accent in Experiment 1*

	Model 1		Model 2		Model 3	
Coefficients	B	SE	B	SE	B	SE
Gender (male)[a]	−0·755**	·245	−0·836**	·257	−0·841**	·258
Homework compliance	−0·001	·013	−0·003	·014	−0·003	·014
Pre-test	4·102***	·444	3·669***	·473	3·667***	·474
Enrichment condition (Eng)[b]			−0·788**	·248	−0·740*	·316
Module (2)[c]			0·617**	·189	0·557*	·262
Module (3)[c]			0·647***	·001	0·795**	·270
Enrichment condition (Eng)[b] × Module (2)[c]					0·119	·351
Enrichment condition (Eng)[b] × Module (3)[c]					−0·276	·353

NOTES: * $p < ·05$; ** $p < ·01$; *** $p < ·001$

[a] Reference group for gender is female.

[b] Reference group for enrichment condition is Putonghua.

[c] Reference group for module is 1.

condition and module was further incorporated into Model 3.[3] Statistical analyses were performed with the R v2.15.1 statistical programming language (R Core Team, 2012), using the *clmm* function in the ordinal package (Christensen, 2012) for R.

Tables 2 and 3 show the model estimates of the multilevel ordinal logistic regression analyses for Putonghua and English accent ratings respectively. For both target languages, only a linear trend in the relation between pre-test and post-test ratings was found to be significant ($ps < ·001$). In the regression analyses for Putonghua accent, only one control variable, namely gender, was a significant predictor of posttest accent ratings ($\beta = -0·841$, p −.001). Specifically, the odds of being rated at a higher accent category decreased by 57% for boys as compared to girls (odds ratio, $OR = ·43$).

Most importantly, the coefficient for enrichment condition was significant ($\beta = -0·740$, $p = ·02$); children in the Putonghua enrichment group scored significantly higher on Putonghua accent than those in the English enrichment group at post-test. Relative to the Putonghua group, the English enrichment group showed a 52% decrease in odds in moving to the next higher level of post-test Putonghua accent rating ($OR = ·48$). Consistent with this, a likelihood-ratio test comparing the models with

[3] The function call for Model 3 was study1.chn.fm2 <- clmm(posttest ~ gender + hw_compliance_percent + pretest + group_story * module + (1|subject_no), data = na.omit (study1.chn))

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

TABLE 3. *Model estimates for the multilevel ordinal logistic regression analyses on English accent in Experiment 1*

	Model 1		Model 2		Model 3	
Coefficients	B	SE	B	SE	B	SE
Gender (male)[a]	−0·297	·162	−0·366	·190	−0·370	·192
Homework compliance	0·020*	·009	0·025*	·011	0·025*	·011
Pre-test	5·496***	·731	4·826***	·776	4·793***	·776
Enrichment condition (Eng)[b]			0·240	·186	−0·017	·259
Module (2)[c]			0·206	·173	−0·158	·233
Module (3)[c]			1·446***	·190	1·436***	·246
Enrichment condition (Eng)[b] × Module (2)[c]					0·722*	·312
Enrichment condition (Eng)[b] × Module (3)[c]					0·040	·316

NOTES: * p <.05; ** p <·01; *** p <·001.
[a] Reference group for gender is female.
[b] Reference group for enrichment condition is Putonghua.
[c] Reference group for module is 1.

and without the enrichment condition as a predictor (i.e. Model 2 and Model 1 respectively) while controlling for all the remaining variables was also significant ($\chi^2(1) = 9{\cdot}93$, $p = {\cdot}002$). Together, these results revealed that children who had received language enrichment in Putonghua showed significant benefits in post-test Putonghua accent scores when compared to those who had received enrichment in English.

Significant main effects of module indicated that the accent scores improved across the three semesters. Importantly, there were no significant interactions between enrichment condition and module (ps > ·43), suggesting that the enrichment benefits generalized well across the three enrichment modules.

By contrast, enrichment condition was not a significant predictor of post-test accent ratings in English ($p = {\cdot}95$). Nonetheless, the interaction between enrichment condition and Module 2 was found to be significant ($\beta = 0{\cdot}722$, $p = {\cdot}02$); the patterns of pre-test and post-test English accent scores in Module 2 hinted at benefits of the English enrichment. For instance, the percentage of children scoring 2·0 or better went up from 32% (pre-test) to 47% (post-test) in the English enrichment group, whereas the percentage remained the same (36% for pre-test as well as post-test) in the Putonghua enrichment group. The percentage of children scoring 3·0 or better went up from 10% to 17% in the English enrichment group, and went from 8% to only 10% in the Putonghua enrichment group. Together, these findings suggested that, unlike its Putonghua counterpart, the English enrichment in this experiment did not yield statistically significant benefits for the

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

children across modules, but there may be some – albeit much less robust – benefits in Module 2.

To summarize, Experiment 1 revealed that occasional non-interactive input from native speakers (e.g. listening to audio-recorded stories read by multiple speakers) helped young children acquire a better accent in learning a second language similar to their first language. There is a hint that such input might also help children acquire a better accent in learning a second language very different from their native language, but the effects – if any – seemed far less robust.

Aside from the issue of target language, a more general question about the robustness of enrichment program benefits remains. Will the enrichment benefits hold up well if program characteristics are altered? Experiment 2 had three main goals. First, it explored the issue of pacing. The enrichment program in Experiment 1 spanned almost 18 months – with stories presented at a leisurely rate of two or three per month – in part because spaced (or distributed) presentation had been found more effective than massed presentation (Ebbinghaus, 1885; see Greene, 2008, for a review). One concern is that during these 18 months, children could have significant exposure to the second languages by native- or near-native speakers outside of school, which could influence the children's phonology development. Parental reports of home language were not collected but would have been informative.

Note, however, that the children were individually and randomly assigned to the two experimental conditions (Putonghua and English). With a relatively large sample such as the present one ($N = 155$), this should have helped to ensure that, on average, the children in the two conditions were comparable in all aspects, including language exposure outside school. It is therefore reasonable to attribute any measurable outcome differences between the conditions to the experimental manipulation. That said, it would be informative to see whether the findings of Experiment 1 are replicable with a shorter intervention period. Experiment 2 sped things up, offering children audios of several stories a week for 7 weeks (compared to two or three stories per month for 18 months in Experiment 1). Would this faster paced program work?

Second, we wanted to see if a language enrichment program with less adult involvement could also be effective. In Experiment 2, we therefore used the same at-home listening enrichment program as in Experiment 1, except that the children's teachers did not introduce the stories at school first.

The third goal was to determine whether any language enrichment benefits detected in Experiment 1 would generalize to novel materials. In Experiment 1, the test stories were selected from the enrichment stories. In Experiment 2, a test story that the children had never seen or heard before was added in each language.

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

EXPERIMENT 2

METHOD

Participants

One hundred and twelve children participated with written parental consent (55 boys and 57 girls; aged from 6;9 to 7;9 at pre-test). The children were native speakers of Cantonese recruited from five second-grade classrooms in Hong Kong (participation rate: 67%). As in Experiment 1, none of these children's parents were native speakers of either English or Putonghua, according to their home-room teachers. The children attended the same public school as the children in Experiment 1, but were two years behind them.

Stimulus materials

The listening-enrichment materials (storybooks and story audios on CDs) and the listening-log worksheets were the same as those in Experiment 1. However, only thirty-four of the thirty-seven stories in each language were used here.

Procedure

As in Experiment 1, children were randomly assigned individually to either the Putonghua or the English enrichment condition, with fifty-six children per condition (Putonghua: 27 girls, 29 boys; English: 30 girls, 26 boys). The take-home listening program took place in October and November of second grade. To offer a faster paced enrichment program, children were asked to listen to four stories in a 'Practice Week' and then five stories per week for another six weeks. Altogether, each child was given thirty-four stories over seven weeks.

The language enrichment program

Unlike in Experiment 1, children in this experiment were simply given only take-home listening assignments. The teachers did not introduce the stories in class. The procedure for the take-home enrichment program was identical to that of Experiment 1, except for the pace of the program.

Word-reading assessment at baseline

The procedure was identical to that of Experiment 1.

Accent assessment

Each child was asked to read four stories aloud, two in each language. For each story, a child was given 2 or 3 minutes to look at the illustrated

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

storybook and then asked to read the story aloud, when the reading was audio-recorded for accent rating. Unlike in Experiment 1, children did not receive an audio demonstration of how to read the test stories; our pilot data suggested that this was not necessary because these second-graders had better Putonghua and English reading abilities than their first-grade counterparts at the first pre-test in Experiment 1. The entire assessment took about 15 minutes. Two of the test stories were randomly selected from the test stories used in Experiment 1: the Putonghua test story used in Module 1 and the English test story used in Module 2 (i.e. enrichment stories used as test stories). In addition, two novel stories from the same storybook series (i.e. *The Magic Box* readers; again one in each language) were used to assess how well intervention effects might generalize to language materials not already used in the enrichment program (i.e. generalization stories). The pre-test and post-test procedure was otherwise similar to that of Experiment 1.

Accent ratings

Two native English speakers independently rated accents on the English test stories, and two native Putonghua speakers independently rated accents on the Putonghua test stories. The accent rating scales and procedure were the same as those of Experiment 1.

RESULTS AND DISCUSSION

Intervention compliance

The percentage of correct responses for thirty-four listening-log worksheets (i.e. 204 items) was 78·7% ($SD = 20·5\%$) in the English enrichment condition and 70·3% ($SD = 20·7\%$) in the Putonghua enrichment condition. The English enrichment program showed slightly better results favoring the English enrichment program, although this difference was not significant ($t(113) = 1·9$, $p > ·06$). In this much faster paced program, the intervention compliance rate – while still quite good – was understandably lower than that of Experiment 1.

Word-reading abilities at baseline

On average, in the Putonghua pre-test, children in the Putonghua and English enrichment conditions correctly read aloud 12·5 words ($SD = 9·2$) and 11·3 words ($SD = 8·6$) respectively. The corresponding averages for the English pretest were 3·7 words ($SD = 4·1$) and 2·9 words ($SD = 3·7$). The two conditions did not differ significantly ($t(110) = 0·70$, $p > .4$ for Putonghua; $t(110) = 1·16$, $p > ·2$ for English).

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

TABLE 4. *Children's mean pre-test and post-test scores (standard deviations) of Putonghua and English accent in Experiment 2*

Putonghua enrichment condition

	Story accent ($N = 56$)	
	Putonghua (with enrichment)	English (without enrichment)
Pre-test	3·36 (0·90)	1·86 (0·59)
Post-test	3·65 (0·71)	2·00 (0·55)

English enrichment condition

	Story accent ($N = 56$)	
	English (with enrichment)	Putonghua (without enrichment)
Pre-test	1·74 (0·61)	3·34 (0·87)
Post-test	1·98 (0·59)	3·49 (0·76)

Inter-rater reliability

The two independent raters for the two Putonghua test stories achieved acceptable inter-rater reliability on the 5-point accent rating scale (Spearman's rhos (233) ranged from ·57 to ·65, $ps < ·001$). Using the rating scale that allowed half-points, the two independent raters for both English test stories also achieved good reliability (Spearman's rhos (233) ranged from ·74 to ·75, $ps < ·001$).

Language enrichment program effects

Table 4 presents the mean pre-test and post-test accent scores for the two enrichment conditions to offer a general idea of the children's Putonghua and English accents. As in Experiment 1, three multilevel ordinal logistic regression models were constructed for each target language. The baseline model (Model 1) examined the fixed effects of the control variables, which included gender, homework compliance, and pre-test accent rating. In Model 2, test story type (enrichment story vs. generalization story) and enrichment condition were added. In Model 3, the interaction between enrichment condition and test story type was further added.

Results of the ordinal logistic regressions for Putonghua and English accent ratings in Experiment 2 are presented in Table 5 and 6 respectively. As in Experiment 1, orthogonal polynomial contrast analyses revealed a significant linear trend ($ps < ·001$) in the association between pre-test and

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

TABLE 5. *Model estimates for the multilevel ordinal logistic regression analyses on Putonghua accent in Experiment 2*

	Model 1		Model 2		Model 3	
Coefficients	B	SE	B	SE	B	SE
Gender (male)[a]	−0·040	·213	−0·034	·211	−0·029	·211
Homework compliance	0·001	·004	0·002	·004	0·002	·004
Pre-test	4·778***	·414	4·846***	·412	4·903***	·415
Enrichment condition (Eng)[b]			−0·409	·210	−0·715*	·281
Test story type (generalization story)[c]			0·304	·182	−0·005	·261
Enrichment condition[b] × test story type[c]					0·604	·367

NOTES: * $p<·05$; ** $p<·01$; *** $p<·001$.
[a] Reference group for gender is female.
[b] Reference group for enrichment condition is Putonghua.
[c] Reference group for test story type is enrichment story.

post-test scores for both target languages. Among all the other predictors, only enrichment condition was significantly related to post-test ratings of Putonghua accent ($\beta=-0·715$, $p=·01$), whereas neither the main effect of test story type ($\beta=-0·005$, $p=·99$) nor the interaction between enrichment condition and story type ($\beta=0·604$, $p=·10$) were significant. These results documented beneficial effects of the Putonghua enrichment program for helping children acquire a better accent in Putonghua, and such improvement generalized across story type regardless of whether the test story had been used as one of the enrichment stories during intervention.

However, no such effects were found for the English enrichment program. Based on the regression analyses, enrichment condition was not a significant predictor of post-test accent ratings in English ($\beta=-0·044$, $p=·90$). That is, the English enrichment did not significantly increase the likelihood of obtaining higher scores on English accent at post-test, when compared to the Putonghua enrichment. The main effect of test story type reflects that one test story elicited higher accent ratings than the other test story. This effect did not interact with enrichment effect.

Intervention compliance and effects

Intervention compliance did not significantly predict post-test accent scores in Experiment 1, which is not surprising perhaps given the near ceiling effects for intervention compliance (over 90% in both intervention groups). In Experiment 2, the compliance rates went down to around 70% for the Putonghua enrichment condition and 79% for the English enrichment

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

TABLE 6. *Model estimates for the multilevel ordinal logistic regression analyses on English accent in Experiment 2*

	Model 1		Model 2		Model 3	
Coefficients	B	SE	B	SE	B	SE
Gender (male)[a]	−0·161	·267	−0·183	·291	−0·182	·295
Homework compliance	0·002	·006	0·001	·006	0·001	·007
Pre-test	6·896***	1·131	6·279***	1·171	6·381***	1·236
Enrichment condition (Eng)[b]			0·170	·293	−0·044	·354
Test story type (generalization story)[c]			1·019***	·199	0·812**	·283
Enrichment condition[b] × test story type[c]					0·408	·367

NOTES: * $p < ·05$; ** $p < ·01$; *** $p < ·001$.
[a] Reference group for gender is female.
[b] Reference group for enrichment condition is Putonghua.
[c] Reference group for test story type is enrichment story.

condition. Still, as in Experiment 1, there were no significant intervention compliance effects (see homework compliance in Tables 5 and 6).

To sum up, Experiment 2 replicated a main finding of Experiment 1, namely, six- and seven-year-olds benefited from non-interactive language input in acquiring a better accent in a second language (i.e. Putonghua) similar to their native Cantonese. Furthermore, this experiment also documented that the enrichment benefits could generalize to novel contents. (Children could read an entirely novel story with a better accent.) Experiment 2 also demonstrated the effectiveness of an enrichment program that was faster-paced than the one in Experiment 1 (7 weeks instead of 18 months) and required less adult involvement – and might therefore be easier to implement. Unlike in Experiment 1, the teachers in Experiment 2 simply handed out and collected take-home listening assignments; they did not have to introduce the stories in class prior to giving children the take-home listening assignments.

By contrast, when the second language (i.e. English) was very different from their native Cantonese, the benefits of such non-interactive language input for acquiring a better accent, if any, seemed far less robust. While the pattern of means in both experiments was consistent with the predicted benefits (Tables 1 and 4), there was only a hint of the predicted English accent improvement in Module 2 of Experiment 1.

GENERAL DISCUSSION

In the era of globalization, it is more important than ever to speak more than one language well, and that makes language education more important too.

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

A major challenge in second-language acquisition is to master its phonology. Actually, it is easy enough for learners who are very young (under age six or seven) and get regular input from native speakers. But when the second language is not the societal language, those conditions are rarely met. The language may be taught in school, but often not until children are older. Even if classes begin at young ages, the teachers are generally not native speakers. Indeed, many second-language learners have had little or no exposure to native speakers during their early years. So it is understandable that many who speak a second language only do so with a noticeable accent that can interfere with communication. Starting second-language education early, hiring more native-speaking teachers, and enhancing teacher training can help. But due to resource constraints, few schools can fully implement these strategies. A potentially cost-effective alternative is to use audio-recorded input from native speakers. Indeed, audios are already widely used in second-language education all over the world. But while this is assumed to be useful, rigorous evaluation of their effectiveness is very much needed.

The present experiments began this process with children at the high end of the hypothesized sensitive period of phonology acquisition. Prior research has suggested that recorded input does not help infants discriminate phonemes in a foreign language, although live social interaction of similar quantity does (Kuhl *et al.*, 2003; cf. Maye *et al.*, 2002). But during their early school years, children become more experienced with TV, videos, and audios of songs, just as they also make considerable gains in attention control and self-regulation. With these advances (Anderson & Hanson, 2010), they would seem to stand a good chance of benefiting from even a modest amount of non-interactive language input.

In both Experiments 1 and 2, children listened to a variety of stories audio-recorded by multiple native speakers. Modeled on the proven principle of high-variability perceptual training in phonology (Bradlow *et al.*, 1999), this program was user-friendly and well received by young children, as indicated by the high level of intervention compliance among first- and second-graders. We experimentally varied the non-interactive language input to sort out its possible causal impact on second-language accent acquisition; obtained unbiased assessment of accent raters blind to the enrichment condition and pre-test/post-test status of individual story recordings; and checked intervention compliance objectively with listening logs. Together, the two experiments suggest that occasional non-interactive language input from native speakers can help young second-language learners acquire a better accent in a second language.

Future research would do well to extend the present findings to younger as well as older language learners. Note that age five to seven (the age range of the children in these two experiments) is probably at the tail end

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

of the sensitive period of phonology acquisition (Flege *et al*., 1999; Oyama, 1976). If the present listening-enrichment program could be adapted for even younger children, more marked benefits might be uncovered.

Adapting this kind of language enrichment program for older learners is also important for two reasons. First, such adaptation, if successful, would benefit precisely those second-language learners who rarely have many good opportunities for live social interaction with native speakers in everyday life. Second, adapting this kind of enrichment for older learners could speak to the theoretical debate on what underlies the young-over-old advantage in second-language phonology and morphosyntax acquisition. As noted earlier, these age effects could be explained by younger learners having greater sensitivity to linguistic input (Kuhl, 2010; Werker & Tees, 2005), but also by their having greater access to good second-language input (Bialystok & Hakuta, 1994, 1999; Flege *et al*., 2006). If comparable enrichment programs could be offered to second-language learners of a wide age range, it would be possible to see whether older learners – who might be shy about talking to native speakers – can benefit as much as young learners from non-interactive input from native speakers. If so, then the 'access' hypothesis will receive some support. If not, then the young-over-old advantage observed in naturalistic L2 research may indeed primarily reflect a sensitive period of language acquisition.

The present two experiments are also relevant to an important issue about the impact of language enrichment on second-language acquisition. Native Cantonese-speaking children benefited from Putonghua enrichment and perhaps also English enrichment, although the robustness of enrichment benefits seems to depend substantially on the target language.

Note that the two target languages in the present experiments differ in at least four potentially important ways. First, Cantonese is much more similar to Putonghua than to English. For instance, unlike English, neither Cantonese nor Putonghua has voiced stops /b, d, g/ in word-final positions (e.g. Flege & Wang, 1989; the 'g' in Putonghua Romanization actually maps onto the unaspirated /k/). As also noted earlier, Cantonese and Putonghua have many grammatical and phonotactic similarities, lexical cognates, and phonological mappings (Ramsey, 1987). Second, English and Putonghua are very different, irrespective of their relation to Cantonese. One is an Indo-European language, and the other is a Sino-Tibetan language. Also, Putonghua phonology, and the phonology of most Chinese dialects, may be intrinsically easier to acquire than English phonology (So & Dodd, 1995).

Third, in our experiments, the children's abilities in these two languages also differed. Their Putonghua was more advanced than their English from the outset. For instance, their Putonghua accent score was about 1·5 points higher than their English accent score on average (Tables 1 and 4).

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

This may have reflected the relative ease of acquiring the Putonghua phonology compared to the English phonology, but it may also have reflected the native or near-native input of Putonghua in the community. For example, although the teachers and parents were not native speakers of the target languages, it is possible that their Putonghua accents were more native-like than their English accents. Perhaps these learning conditions enabled the children to make better use of the Putonghua enrichment than of the English enrichment, leading to much more robust enrichment program benefits for Putonghua than for English.

Fourth, the English enrichment stories were recorded by native speakers of American-English who had lived both in and outside the US, whereas the children's English teachers spoke the Hong Kong dialect influenced by both British English and Chinese phonology. Such dialectic differences may have made it harder for the children to benefit from the audio-recorded English enrichment stories.

We are mindful that different accent-rating scales were used for the two languages. One concern is that increases in the Putonghua condition for Putonghua accent scores might naturally be quantitatively larger (in full points) than those in the English condition for English accent scores (in half-points). But this does not have to be so. Note that accent ratings can be rounded up as well as rounded down. For example, an English accent rater would probably give a '3' to an accent really worth 2·6 and a '2·5' to an accent really worth 2·4. A Putonghua rater would probably give a '3' and a '2', respectively, instead. With the relatively large number of accent ratings in each language, the full-point scale (for Putonghua) does not necessarily mean larger increases than the half-point scale (for English).

We are also mindful that, even with the listening logs, we could only know that children had listened to the audio-recorded stories at least once. We could not tell how many times the children actually listened and whether they read along out loud. As a first step, these two experiments demonstrated the effectiveness of our enrichment programs. What factors may mediate or moderate the effectiveness of a program should be explored in future research.

In both Experiments 1 and 2, the enrichment benefits were statistically significant for Putonghua but not – except for Module 2 in Experiment 1 – for English, although the pattern of means was consistently in the predicted direction for both target languages. The differences between the two target languages, as just described, may account for some or all of the difference in their respective enrichment benefits, but that has yet to be determined.

Nonetheless, occasional non-interactive language input – as instantiated in the audio storybooks enrichment program – turns out to be an effective educational tool. Just listening to thirty or so very short stories, each read aloud by several native speakers, yielded measurable benefits in the

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

children's second-language accents. For Putonghua, at least, benefits were found when the stories were spread over 18 months as well as when they were concentrated in 7 weeks; benefits were found with and without adults introducing the stories to the children to prepare them for the listening enrichment; benefits generalized well beyond the enrichment materials to novel materials. Moreover, at least in Experiment 1, there was a hint of benefits for languages as diverse as English and Putonghua. Such non-interactive input from native speakers may well prove to be a cost-effective strategy to enrich the language environment of young second-language learners. Beyond these applied implications, the present findings also offer more evidence for the causal relation between perception and production in language acquisition: Merely hearing a language can improve the perception and hence the mental model of its phonology, which can, in turn, guide and improve production (Au, 2013; Best, 1994; Kuhl & Meltzoff, 1982).

REFERENCES

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science* **8**, 199–209.

Anderson, D. R. & Hanson, K. G. (2010). From blooming, buzzing confusion to media literacy: the early development of television viewing. *Developmental Review* **30**, 239–55.

Au, T. K. (2013). Songs as ambient language input in phonology acquisition. *Language Learning and Development* **9**, 266–77.

Au, T. K., Knightly, L. M., Jun, S.-A. & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science* **13**, 238–43.

Au, T. K., Oh, J. S., Knightly, L. M., Jun, S.-A. & Romo, L. F. (2008). Salvaging a childhood language. *Journal of Memory & Language* **58**, 998–1011.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (eds), *Joint attention: its origins and role in development*, 131–58. Hillsdale, NJ: Erlbaum.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (eds), *The development of speech perception: the transition from speech sounds to spoken words*, 167–224. Cambridge, MA: MIT Press.

Bialystok, E. & Hakuta, K. (1994). *In other words: the science and psychology of second-language acquisition*. New York: Basic Books.

Bialystok, E. & Hakuta, K. (1999). Confounded age: linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (ed.), *Second language acquisition and the Critical Period Hypothesis*, 161–181. Mahwah, NJ: Erlbaum.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language* **68**, 706–55.

Birdsong, D. & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisitions. *Journal of Memory & Language* **44**, 235–49.

Bongaerts, T., van Summeren, C., Planken, B. & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition* **19**, 447–65.

Bowers, J. S., Mattys, S. L. & Gage, S. H. (2009). Preserved implicit knowledge of a forgotten childhood language. *Psychological Science* **20**, 1064–9.

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B. & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: long-term retention of learning in perception and production. *Perception & Psychophysics* **61**, 977–85.
Chan, A. Y. W. (2010). An investigation into Cantonese ESL learners' acquisition of English initial consonant clusters. *Linguistics* **48**, 99–141.
Chiswick, B. R. & Miller, P. W. (2004). Linguistic distance: a quantitative measure of the distance between English and other languages. *IZA Discussion Paper Series*, **No. 1246**, online: <http://hdl.handle.net/10419/20510>.
Christensen, R. H. B. (2012). Ordinal – Regression Models for Ordinal Data R package version 2012·09-11, online: <http://www.cran.r-project.org/package=ordinal/>.
Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language* **25**, 141–68.
Curtiss, S. (1977). *Genie: a psycholinguistic study of a modern-day 'wild child'*. New York: Academic Press.
De Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic productions. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage & J. Morton (eds), *Developmental neurocognition: speech and face processing in the first year of life*, 353–63. Dordrecht: Kluwer.
Dehaene-Lambertz, G. & Houston, D. (1998). Faster orientation latencies toward native language in two-month-old infants. *Language & Speech* **41**, 21–43.
Ebbinghaus, H. (1885/1964). *Memory: a contribution to experimental psychology*, trans. H. A. Ruger & C. E. Bussenius. New York: Dover.
Eimas, P. D., Siqueland, E. R., Jusczyk, P. & Vigorito, J. (1971). Speech perception in infants. *Science* **171**, 303–6.
Ferguson, C. A. (1977). Baby talk as a simplified register. In C. Snow & C. Ferguson (eds), *Talking to children: language input and acquisition*, 209–35. Cambridge and New York: Cambridge University Press.
Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Development* **60**, 1497–510.
Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H. & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics* **34**, 153–75.
Flege, J. E. & Wang, C. (1989). Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t/–/d/ contrast. *Journal of Phonetics* **17**, 299–315.
Flege, J. E., Yeni-Komshian, G. H. & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory & Language* **41**, 78–104.
Greene, R. L. (2008). Repetition and spacing effects. In H. L. Roediger III (ed.), *Learning and memory: a comprehensive reference. Vol. 2: cognitive psychology of memory*, 65–78. Oxford: Elsevier.
Ho, C. S.-H., Chan, D. W.-O., Lee, S.-H., Tsang, S.-M. & Luan, V. H. (2004). Cognitive profiling and preliminary subtyping in Chinese developmental dyslexia. *Cognition* **91**, 43–75.
Hyltenstam, K., Bylund, E., Abrahamsson, N. & Park, H.-S. (2009). Dominant-language replacement: the case of international adoptees. *Bilingualism: Language and Cognition* **12**, 121–40.
Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* **21**, 60–99.
Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
Knightly, L. M., Jun, S.-A., Oh, J. S. & Au, T. K. (2003). Production benefits of childhood overhearing. *Journal of the Acoustical Society of America* **114**, 465–74.
Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron* **67**, 713–27.
Kuhl, P. K. & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science* **218**, 1138–41.

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

Kuhl, P. K. & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: vocal imitation and developmental change. *Journal of the Acoustical Society of America* **100**, 2425–38.
Kuhl, P. K., Tsao, F.-M. & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of National Academy of Science USA* **100**, 9096–101.
Mair, V. H. (1991). What is a Chinese 'dialect/topolet'? Reflections on some key Sino-English linguistic terms. *Sino-Platonic Papers* **29**, online: <http://www.sino-platonic.org/complete/spp029_chinese_dialect.html>.
Mayberry, R. I. (1993). First-language acquisition after childhood differs from second language acquisition: the case of American Sign Language. *Journal of Speech & Hearing Research* **36**, 1258–70.
Maye, J., Werker, J. F. & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82**, B101–11.
Mehler, J. & Christophe, A. (1995). Maturation and learning of language in the first year of life. In M. S. Gazzaniga (ed.), *The cognitive neurosciences: a handbook for the field*, 943–54. Cambridge, MA: MIT Press.
Mezzacappa, E. (2004). Alerting, orienting, and executive attention: developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Development* **75**, 1373–86.
Mischel, W. & Ayduk, O. (2002). Self-regulation in a cognitive-affective personality system: attentional control in the service of the self. *Self and Identity* **1**, 113–20.
Oh, J. S., Au, T. K. & Jun, S.-A. (2010). Early childhood language memory in the speech of perception of international adoptees. *Journal of Child Language* **37**, 1123–32.
Oh, J. S., Jun, S.-A., Knightly, L. M. & Au, T. K. (2003). Holding on to childhood language memory. *Cognition* **86**, B53–64.
Oshima-Takane, Y. (1988). Children learn from speech not addressed to them: the case of personal pronouns. *Journal of Child Language* **15**, 95–108.
Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* **5**, 261–83.
R Core Team (2012). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, online: <http://www.R-project.org/>.
Ramsey, S. R. (1987). *The languages of China*. Princeton, NJ: Princeton University Press.
Rvachew, S. (1994). Speech perception training can facilitate sound production learning. *Journal of Speech & Hearing Research* **37**, 347–57.
Sachs, J., Bard, B. & Johnson, M. L. (1981). Language learning with restricted input: case studies of two hearing children of deaf parents. *Applied Psycholinguistics* **2**, 33–54.
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. *Developmental Science* **14**, 949–59.
Snow, C., Arlman-Rupp, A., Hassing, Y., Jobse, J., Joosten, J. & Vorster, J. (1976). Mothers' speech in three social classes. *Journal of Psycholinguistic Research* **5**, 1–20.
So, C. K. & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language & Speech* **53**, 273–93.
So, L. K. H. & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language* **22**, 473–95.
Tees, R. C. & Werker, J. F. (1984). Perceptual flexibility: maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology* **38**, 579–90.
Tomasello, M. & Farrar, M. J. (1986). Joint attention and early language. *Child Development* **57**, 1454–63.
Todd, P. & Aitchison, J. (1980). Learning language the hard way. *First Language* **1**, 122–40.

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press

Werker, J. F. & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development* **7**, 49–63.
Werker, J. F. & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology* **46**, 233–51.
White, L. & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research* **12**, 233–65.
Wilkinson, G. S. (1993). *Wide Range Achievement Test–Revision 3*. Wilmington, DE: Jastak Association.

https://doi.org/10.1017/S0305000913000627 Published online by Cambridge University Press