Running Head: Ambient Language Input
In press, *Language Learning and Development*

Songs as Ambient Language Input in Phonology Acquisition

Terry Kit-fong Au
The University of Hong Kong

Acknowledgments:
I 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). I am grateful to a team of dedicated research assistants for data collection, coding, and analysis, and to Karen Ravn for her valuable input on an earlier draft. Correspondence Address: Department of Psychology, The University of Hong Kong, Pokfulam Road, Hong Kong, China, or terryau@hku.hk.

Abstract

Children cannot learn to speak a language simply from occasional non-interactive exposure to native speakers' input (e.g., by hearing television dialogues), but can they learn something about its phonology? To answer this question, the present study varied ambient hearing experience for 126 5- to 7-year-old native Cantonese-Chinese speakers learning English as a second language and Putonghua as a second dialect. Hearing Putonghua songs as ambient sounds improved the children's Putonghua accent in a story read-aloud task, whereas hearing English songs yielded no measurable analogous benefit. Research and educational implications are discussed.

Songs as Ambient Language Input in Phonology Acquisition

The prevailing wisdom is that children cannot learn to speak a language simply by having occasional non-interactive exposure to it (Pinker, 1994). Indeed, the hearing children of profoundly deaf parents acquire little oral language just by watching TV frequently with the sound on (Sachs, Bard, & Johnson, 1981; Todd, 1975; Todd & Aitchison, 1980). Similarly, children do not acquire the language of a neighboring country by merely watching TV broadcast from across the border (Snow et al., 1976). But can occasional non-interactive exposure to a language teach children anything about it? Some evidence suggests that preschool children (aged 3 to 5 years) can acquire new vocabulary by watching TV (e.g., Rice & Woodsmall, 1988). The present study tested whether non-interactive exposure to a second language – through audios of songs sung by native speakers – can help young children acquire its phonology.

When exposed to a second language spoken by its native speakers, young children acquire its phonology more readily than older children, who in turn outperform adults (Flege, Yeni-Komshian, & Liu, 1999; Oyama, 1976). This is perhaps because younger learners have more access to good second-language input by virtue of having more interaction with native speakers of the language (Bialystok & Hakuta, 1994, 1999; Flege et al., 2006) and because they are more sensitive to linguistic input in general (Newport, 1990; Werker & Tees, 2005; Kuhl, 2010). Whatever the reason, children's phonological feats begin early. Relying on characteristic intonational and rhythmic patterns, 2-month-olds come to prefer their ambient language to other languages, even when phonetic cues have been removed by low-pass filtering (Dehaene-Lambertz & Houston, 1998; Mehler, Dupoux, Nazzi, & Dehaene-Lambertz, 1996). Preverbal infants also learn to categorize speech sounds into consonants and vowels relevant to their ambient language (Eimas, Siqueland, Jusczyk, Vigorito, 1971; Jusczyk, 1997; Kuhl & Iverson, 1995; Werker & Tees, 1984). Once acquired, such speech sound categories seem to persist for years even after exposure to the childhood language has ended (e.g., due to emigration or change of caregivers; Bowers, Mattys, & Gage, 2009; Oh, Au, & Jun, 2010; Tees & Werker, 1984). However, the amount of exposure does seem to matter.

While infants can learn much about the phonology of an ever-present ambient language, they do not seem to benefit from hearing occasional televised or audio-recorded speech. In a foreign-language intervention experiment, American 9-month-olds learned to contrast Putonghua (also known as Mandarin) Chinese consonants from twelve 25-minute sessions of live social interaction with a native Putonghua speaker, but not from similar quantities of televised or audio-only presentations (Kuhl, Tsao, & Liu, 2003; Kuhl, 2007; but see Maye, Weiss, & Aslin, 2007; Maye, Werker, & Gerken, 2002). Perhaps infants rely on social interaction to heighten their attention to sounds most likely to help them acquire the language of their caregivers. Interaction with adults may serve as a "social gate" to filter out incidental and random sights and sounds to some extent. It remains to be seen whether occasional ambient language input can help older children acquire the phonology of a second language.

University students taking Spanish as a second language learn to speak with a better (i.e., more native-like) accent if they overheard Spanish as children – even if their exposure was for only a few hours a week during large family parties; they spoke English almost exclusively while they were growing up; and their spoken Spanish was limited primarily to

single words and phrases (e.g., names for people, Latino food, places, and festivals). Their advantage in Spanish phonology is substantial and pervasive, holding up well at various levels: individual phoneme (e.g., voice onset time of /b, d, g/ and /p, t, k/), phonological rule (e.g., producing /b, d, g/ as lenited instead of stop consonants in both within-word and across-word intervocalic contexts), whole-sentence accent, and narrative accent (Au, 2012; Au, Knightly, Jun, & Oh, 2002; Au, Oh, Knightly, Jun, & Romo, 2008; Knightly, Jun, Oh, & Au, 2003; Montrul, 2009). But these are correlational findings, and correlation does not mean causation. It may be that merely overhearing a language spoken by native speakers for a few hours a week can help young children acquire its phonology. But that remains unclear.

To see if, and how much, second-language learners benefit from occasional non-interactive exposure to native speakers of the language during childhood, it is necessary to vary their hearing experience experimentally. Hong Kong offers an invaluable window on this question. The ambient language for most Hong Kong Chinese children is Cantonese Chinese. English and Putonghua (also known as Mandarin) are taught in school to most children from age 5 or 6 on, mostly by non-native-speakers for 2 to 3 hours per week. Language input from native speakers is generally minimal outside school (e.g., brief announcements on subway trains); viewing TV in either language is generally rare for young children with limited ability in these languages. This makes it relatively easy to introduce experimental variations in children's exposure to input from native speakers of English and Putonghua.

In this experiment, Cantonese-Chinese first-grade children were randomly assigned to hearing audio-recorded children's songs in either English or Putonghua during lunch at school. Songs sung by native speakers were chosen as enrichment materials for two reasons. First, these first-graders were beginner learners of English and Putonghua as a second language/dialect. Given their limited abilities in these languages, they might find songs more interesting and enjoyable than speech.

Second, formal music training seems to improve phonological processing (e.g., Besson, Chobert, & Marie, 2011; Kraus & Chandrasekaran, 2010; Marie, Delogu, Lampis, Olivetti Belardinelli, & Besson, 2011; Moreno et al., 2009; Slevc & Miyake, 2006). Even short-term music training (an hour per day for 20 days) boosted 4- to 6-year-olds' verbal IQ scores, whereas visual-art training did not (Moreno et al., 2011). A music training helped children extract "words" from a continuous flow of an artificial language, whereas a painting program did not (François,Chobert, Besson, & Schön, 2012). Moreover, entirely passive (i.e., non-social) exposure to foreign music has been found to enhance learning of foreign musical structures (Hannon & Trehub, 2005; Trainor & Corrigall, 2010). Perhaps there is something about music that facilitates learning from passive, non-social exposure?

In the present experiment, Cantonese-Chinese children were randomly assigned to hearing audio-recorded children's songs in either English or Putonghua during lunch-period at school for 18 weeks. Their English and Putonghua accents were assessed with story read-aloud tasks both before and after the language enrichment program. Productive phonology was used to evaluate possible benefits because, in theory, hearing a language during childhood could improve the perception and hence the mental model of its phonology, which could in turn improve production (e.g., Best, 1994; Flege, 1995). The English 'r'-'l' contrast, for instance, eludes many Japanese adults. Nonetheless, perceptual training helps them produce the English 'r' and 'l' better (Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999; Bradlow, Pisoni, Akahane-Yamada & Tohkura, 1997). Perceptual training in speech therapy (e.g., Rvachew, 1994) has likewise demonstrated that improving speech perception can improve production too.

Merely hearing a second language in native speakers' speech or songs, then, may turn out to improve the learners' productive phonology as well. While it would be informative to assess both production and perception, general improvement in production (e.g., global accent) is easier to assess than overall improvement in perception. So as a first step, this experiment assessed productive phonology with accent ratings by native speakers.

In this experiment, the children were native speakers of Cantonese Chinese who were learning both English as a second language and Putonghua as a second dialect. Cantonese and Putonghua are very distinct dialects; native Cantonese-speakers do not understand Putonghua on initial contact, and vice versa for native Putonghua-speakers with Cantonese. Linguistically, they are different enough to be considered different languages (e.g., Ramsey, 1987). However, similar grammars, lexical cognates, and phonological mappings (e.g., /k^h/ and /w/ in Putonghua map onto /h/ and /m/ in Cantonese for many words) render Putonghua easier than English for Cantonese-speaking children to acquire (e.g., Chiswick & Miller, 2004). It remains to be seen whether the impact of hearing songs sung in a second language depends on similarity between that language and a learner's native language.

The present experiment, therefore, will address two research questions:

1. In acquiring the phonology of a second language, can young children benefit from occasional non-interactive exposure to that language, e.g., through songs sung by native speakers?
2. Does such benefit depend on the linguistic relatedness between the learners' native language and the second language?

Methods

Participants

126 first-grade children (55 boys and 71 girls; age ranged from 5;11 to 6;10 at pretest) in a public primary school in Hong Kong participated with written parental consent. They were native speakers of Cantonese Chinese (the ambient language in Hong Kong) and beginners in learning English and Putonghua Chinese as a second language/dialect. Their school offered English and Putonghua classes that were all taught by non-native speakers for about two to three hours per week per language.

Stimulus Materials

The language enrichment materials were audio-recordings of children's songs in English and Putonghua. The songs had simple, repetitive, clearly enunciated lyrics and sparse musical accompaniment. The songs for the English enrichment included the 19 songs collected in a commercially available music CD "Raffi's Singable Songs for the Very Young" sung by a Canadian male singer with guitar accompaniment. But the only commercial recordings we found of children's songs in Putonghua were sung by choirs, which made the lyrics more difficult to discern than recordings of a solo adult vocalist. So the 19 songs for the Putonghua enrichment were recorded specifically for this experiment by an adult male native speaker of Putonghua with piano accompaniment. The titles of songs used in two enrichment program are listed in Appendix A. (Audios of the Putonghua songs are available at http://www.psychology.hku.hk/download/chinese_songs/index.html.) The songs were played on a CD-player with an amplifier in the children's classrooms, where they had their lunch at school.

Procedure

English/Putonghua enrichment program. The children came from four first-grade classrooms, with two randomly assigned to hearing Putonghua songs (65 children; 57% girls), and the other two, to English songs (61 children; 56% girls). The program was administered during five lunch periods per week for 18 weeks. After the children had settled down with their lunch boxes, the teacher in each classroom played the Putonghua/English songs as background music for about 20 minutes on average (ranging from about 15 to 25 minutes).

Because both enrichment groups received identical treatment except for the language of enrichment, they served as each other's tailor-made control. If the enrichment was effective, children randomly assigned to the Putonghua enrichment should have improved more in Putonghua accent than those assigned to the English enrichment. Likewise, children randomly assigned to the English enrichment should have improved more in English accent than those assigned to the Putonghua enrichment.

It took about one week to play all 19 songs in each classroom, so each song was played around 18 times over the 18-week enrichment. According to the teachers, and confirmed by research assistants' observations, the songs were audible throughout each classroom. Some children swayed or tapped to the rhythms of the songs or even sang along; other children chatted with their friends or did homework.

English/Putonghua accent assessments. To see if the enrichment program benefited children's phonology, we audiotaped them individually both before and after the enrichment program as they read a simple story in English and another in Putonghua. The storybooks were selected from two similar series written for beginner readers published in Hong Kong (*The Magic Box* readers available at www.greenfieldhk.com). Each story was 7 pages long with one sentence/phrase per illustrated page. The English story read as follows: *The football game. I can run. I can throw. I can jump. I can kick. I can score the goal. "Hooray!"* The Chinese story could be translated as follows: "Run quickly. Rabbit run quickly. Deer run quickly. Lion run quickly. Tiger run quickly. Zebra run quickly. Monkey run quickly. A big forest fire!" (The text is available upon request.)

The children were tested twice with identical procedures individually in a quiet room at school (pretest and posttest). The pretest was done in the third month of the first grade, when their reading abilities in English and Putonghua Chinese were quite limited. To make the reading task more feasible, a research assistant first showed each child one of the illustrated storybooks and let the child look at it for two or three minutes, then asked the child to listen to an audio of a native speaker reading the story, and finally asked the child to read it aloud and audio-recorded it for accent rating. The procedure was then repeated with the story in the other language. The order of the stories (English vs. Putonghua) was randomized and counterbalanced across children, and the assessment took on average about 8 minutes. The posttest was conducted within two weeks after the enrichment program was completed (near the end of first grade).

Three research assistants did the testing simultaneously in separate rooms. A helper brought children from different classrooms to the testing rooms to minimize the chance that the research assistants might be influenced by knowing children's enrichment conditions.

Accent ratings. The children's audio-recordings were edited to yield one sound file per story-reading. Three adult native speakers of Putonghua and three adult native speakers

of English independently rated how native-like the children's accents were when they read the stories aloud. The Putonghua accent raters grew up in northern Mainland China (1 male, 2 female); they had lived in Hong Kong for several months. One of the English accent raters (female) grew up in Canada and was in Hong for a semester as an exchange student; the other two (1 male, 1 female) grew up in Hong Kong with English as their home language and, since age three or four, had attended international schools staffed predominantly by native-speakers of English. All six raters were either undergraduates or university graduates in psychology without linguistics training. The ratings were done on a five-point 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). Accent ratings of Spanish using this scale correlated well with acoustical measurements at the phonemic level (e.g., voice onset time of /b, d, g/ and /p, t, k/), and coding at the phonological-rule level (e.g., producing /b, d, g/ as lenited instead of stop consonants in within-word and across-word intervocalic contexts), with *r*s ranging from .5 to .78 (Au et al., 2002; Knightly et al., 2003).

Each accent rater listened to all 252 story-readings produced by the 126 children at both pretest and posttest. The story recordings were presented in a random order on a computer with the E-Prime software and headphones. So the raters did not know whether a specific story-reading had been produced by a child who had heard Putonghua songs or English songs or whether it had been produced at pretest or posttest. Each rater independently listened to each story recording and then entered an accent score into the computer. The stories were presented in blocks of 40 stories, evenly divided between pretest and posttest and between the two language enrichment conditions in each block. The ratings were self-paced and took about six hours to complete.

Raters were asked to attend to phonological characteristics of English and Putonghua that are especially challenging for native Cantonese speakers (Chan, 2010; Flege & Wang, 1989; So & Best, 2010): prosody (e.g., intonation contours, stress patterns of multi-syllabic words), lexical tones for Putonghua (e.g., the dipping tone), and phonemes (e.g., the English ‘r’, word-final /t/ and /d/, voicing of /b, d, g/; the Putonghua alveolar /z, c, s/ and retroflex /zh,ch, sh/, /r/ consonants).

Results

The Putonghua accent scores ranged from 1 to 5 on the five-point rating scale, and the inter-rater reliability on the full data set of 252 stories among the three raters was excellent (intraclass correlation $R = .91$). By contrast, the accent scores for the English story clustered at the low end, with mostly 1 and 2 and an occasional 3. Several factors may have contributed to better accent scores in Putonghua than in English. First, Putonghua phonology may be easier for Cantonese-Chinese children because it is more similar to Cantonese than is English. Second, the children’s Putonghua teachers might have had more native-like accents than the English teachers, although all were non-native speakers. Third, the children may have received more exposure to Putonghua than English through the media in Hong Kong. Fourth, the raters of the children’s English accents may have used more stringent criteria than the raters of Putonghua accents.

Regardless of why the children received better accent scores in Putonghua than in English, a more discriminating scale was needed for rating their English accents. Half-points were therefore used for the English accent ratings (i.e., 1, 1.5, 2, 2.5, 3,...). Among the three

raters, the inter-rater reliability on the 252 English stories was high for this revised scale (intraclass $R = .79$).

In the Putonghua Enrichment condition, children's mean Putonghua accent score went up from 2.68 (SD = .86) on the pretest to 3.24 (SD = .70) on the posttest. In the English Enrichment condition, the mean Putonghua accent score went up from 2.83 (SD = .83) on the pretest to 3.07 (SD = .82) on the posttest. An improvement score was computed for each child in each language by subtracting the pretest score from the posttest score. An independent-sample t-test revealed that children's Putonghua accent improvement in the Putonghua Enrichment condition was significantly larger than that in the English Enrichment condition ($t(124) = 2.73$, $p < .01$, 2-tailed, Cohen's $d = .49$; Cohen, 1977; see Figure 1).

The average improvement in Putonghua accent from pretest to posttest for children who had heard Putonghua songs (Mean = .56, SD = .66) had a rather large effect size (Cohen's $d = .85$; one-sample $t(64) = 6.8$, $p < .001$, two-tailed). By contrast, the average improvement in Putonghua accent for children who had heard English songs instead (Mean = .24, SD = .65) had only a small effect size (Cohen's $d = .37$; one-sample $t(60) = 2.9$, $p < .01$, two-tailed).

Analogous analyses were conducted on the English accent scores. In the English Enrichment condition, children's mean English accent score went up from 1.35 (SD = .47) on the pretest to 1.61 (SD = .39) on the posttest. In the Putonghua Enrichment condition, the mean English accent score went up from 1.29 (SD = .41) on the pretest to 1.58 (SD = .42) on the posttest. An independent-sample t-test was conducted on the improvement scores for English accents. No significant difference was found between the two enrichment conditions ($t(124) = .61$, $p > .5$, 2-tailed, Cohen's $d = .09$; see Figure 1). Nonetheless, children's accent scores improved from pretest to posttest in both the English enrichment condition (Mean = .26, SD = .37; one-sample $t(60) = 5.5$, $p < .001$, two-tailed; Cohen's $d = .70$) and the Putonghua enrichment condition (Mean = .29, SD = .29; one-sample $t(64) = 8.0$, $p < .001$, two-tailed; Cohen's $d = 1.0$). Children's accents in both languages generally improved from pretest to posttest. Such improvement may be due to several months of English and Putonghua classes in first grade and/or due to a practice effect from reading the same test stories aloud twice. In any case, the accent ratings picked up that improvement, suggesting that the accent ratings were reasonably sensitive. Importantly, enrichment benefits were found for only Putonghua songs and not English songs.

Discussion

This language enrichment experiment suggests that Cantonese-speaking children learning Putonghua as a second dialect can benefit from hearing songs sung by a native Putonghua speaker. Previous research has documented a correlation between overhearing live conversations in a second language during childhood and later learning to speak it with a more native-like accent. The present experiment suggests that occasional ambient indirect input – even if only from song audios rather live speakers – can help young children acquire a better phonology. These findings are in line with those of longitudinal studies demonstrating that music training impacts on several aspects of children's language processing, such as the speech segmentation of an artificial language (François et al., 2012) and the reading of phonological complex words (Moreno et al., 2009; see also Kraus & Chandrachekaran, 2010).

It remains to be seen how well the present findings generalize to second-language phonology acquisition. For the young native Cantonese speakers in the present experiment, Putonghua was a second dialect, rather than a second language. But as noted earlier, these two dialects are not mutually intelligible (Mair, 1991) and are considered by some linguists to be different languages (e.g., Ramsey, 1987). They are probably no more similar than two closely related languages such as Spanish and Portuguese (Chiswick & Miller, 2004). Still, it remains to be seen whether this kind of language enrichment will, indeed, benefit children learning a second language closely related to their first language.

No measurable benefits of such enrichment were detected for children learning a second language (English) not closely related to their first language (Cantonese). There are several plausible reasons why benefits were detected for Cantonese but not English. First, the children's Putonghua was more advanced than their English from the outset; on average their accent scores were about 1.5 points higher in Putonghua than in English. Perhaps this kind of language exposure is only useful after children have some threshold knowledge of the second language. Second, the dialectic differences between the English enrichment songs (i.e., Raffi's songs in a North American dialect of English) and the English spoken by the children's English teacher (i.e., the Hong Kong dialect influenced by both British English and Chinese phonology) may have made it harder for the children to benefit from the English enrichment songs.

Third, different test stories were used for assessing the children's Putonghua accent and English accent, although the storybooks were selected from two similar series written for beginner readers published in Hong Kong. In future research, using the same story published in both Chinese and English can help rule out the possibility that the text might be intrinsically more challenging in one language than the other. Fourth, the accent ratings might not be sensitive enough to detect subtle improvement in accent for beginner learners. Recall that children's Putonghua accent scores ranged from 1 to 5, whereas their English accent scores clustered at the low end, with mostly 1 and 2 and an occasional 3. Although good inter-rater reliability was achieved for English using the finer scale with half-points, the accent assessment still might not be sensitive enough.

Research with older children might have different results. Second- or third-graders' English abilities might be more comparable to the Putonghua abilities observed in first-graders in the present study, and hearing English songs may turn out to have measureable benefits for them.

One limitation of the present experiment is its focus on only productive phonology. It remains to be seen whether enrichment benefits can also be found in perception. The present set of enrichment songs is probably too diffuse to improve perception of specific phonological characteristics that can be readily evaluated. However, more strategically selected songs could target phonemic or lexical tone contrasts especially challenging for the second-language learners. Such research can yield a fuller understanding of the potential benefits of songs as input in learning second-language phonology. It will also be informative in future research to compare children receiving an enrichment program with those who do not. Such a comparison can help tease apart the effect of maturation and regular language education from the effect of enrichment.

The present findings have several applied implications. Audios and videos are commonly used in second-language education. Rigorous evaluation of their effectiveness can

help parents and teachers make informed decisions about how best to use young second-language learners' time. This is an especially important question when a second language is not the societal language. In that case, live interaction with native speakers is often limited for children during their preschool or early school years, although that may be when such input would be most useful for acquiring the target phonology (Flege, Yeni-Komshian, & Liu, 1999; Oyama, 1976). Even when second-language education begins early, teachers are generally not native speakers of the target language. It is no wonder that many second-language learners – in such learning contexts – speak with a noticeable accent that could interfere with communication. How can we do better? Hiring more native-speaking teachers and enhancing teacher training seem like sensible options. But few schools can fully implement these strategies due to resource constraints. One potentially cost-effective approach might be to use audio-recorded input from native speakers. The present language enrichment experiment suggests that this could work.

References

Au, T. K. (2012). Access to Childhood Language Memory: Implications for Cognitive Development. In M. Siegal & L. Surian (Eds.), *Access to language and cognitive development* (pp. 176-191). New York: Oxford University Press.

Au, T. K., Knightly, L. M., Jun, S.-A. & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science, 13*, 238-243.

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.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology, 2*, 1-12.

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* (pp. 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: Second language acquisition research* (pp. 161-181). Mahwah, NJ: Erlbaum.

Bowers, J. S., Mattys, S. L., Gage, S. H. (2009). Preserved implicit knowledge of a forgotten childhood language. *Psychological Science, 20*, 1064-1069.

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*, 977-985.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R. & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101*, 2299-2310.

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,* http://hdl.handle.net/10419/20510.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Dehaene-Lambertz, G. & Houston, D. (1998). Faster orientation latencies toward native language in two-month-old infants. *Language & Speech, 41*, 21-43.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303-306.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 233-272). Baltimore: York.

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*, 153-175.

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.

François, C., Chobert, J., Besson, M., & Schön, D. (2012). Music training for the development speech segmentation. *Cerebral Cortex,* published online July 10, 2012.

Hannon, E. E., & Trehub, S. E. (2005). Turning in to musical rhythms: Infant learn more readily than adults. *Proceedings of the National Academy of Sciences, 103,* 12639-12643.

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,* 465-474.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Review Neuroscience, 11*, 599-605.

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science, 10*, 110–120.

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron, 67*, 713-727.

Kuhl, P. K. & Iverson, P. (1995). Linguistic experience and the 'perceptual magnet effect.' In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 121-154). Baltimore: York.

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,* 9096–9101.

Mair, V. H. (1991). What is a Chinese "dialect/topolet"? Reflections on some key Sino-English linguistic terms. *Sino-Platonic Papers, 29,* http://www.sino-platonic.org/complete/spp029_chinese_dialect.html

Marie C., Delogu F., Lampis G., Olivetti Belardinelli M., Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience, 23*, 2701-2715.

Maye, J., Weiss, D. J., & Aslin, R. N. (2007). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11*, 122-134.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101-B111.

Mehler, J., Dupoux, E., Nazzi, T. & Dehaene-Lambertz, G. (1996). Coping with linguistic diversity: The infant's viewpoint. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 101-116). Mahwah, NJ: Erlbaum.

Montrul, S. (2009). Reexamining the fundamental difference hypothesis: What can early bilinguals tell us? *Studies in Second Language Acquisition, 31*, 225-257.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E.G., Cepeda, N.J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science, 22*, 1425-1433.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S.L., & Besson, M. (2009). Musical Training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cerebral Cortex, 19*, 712-723.

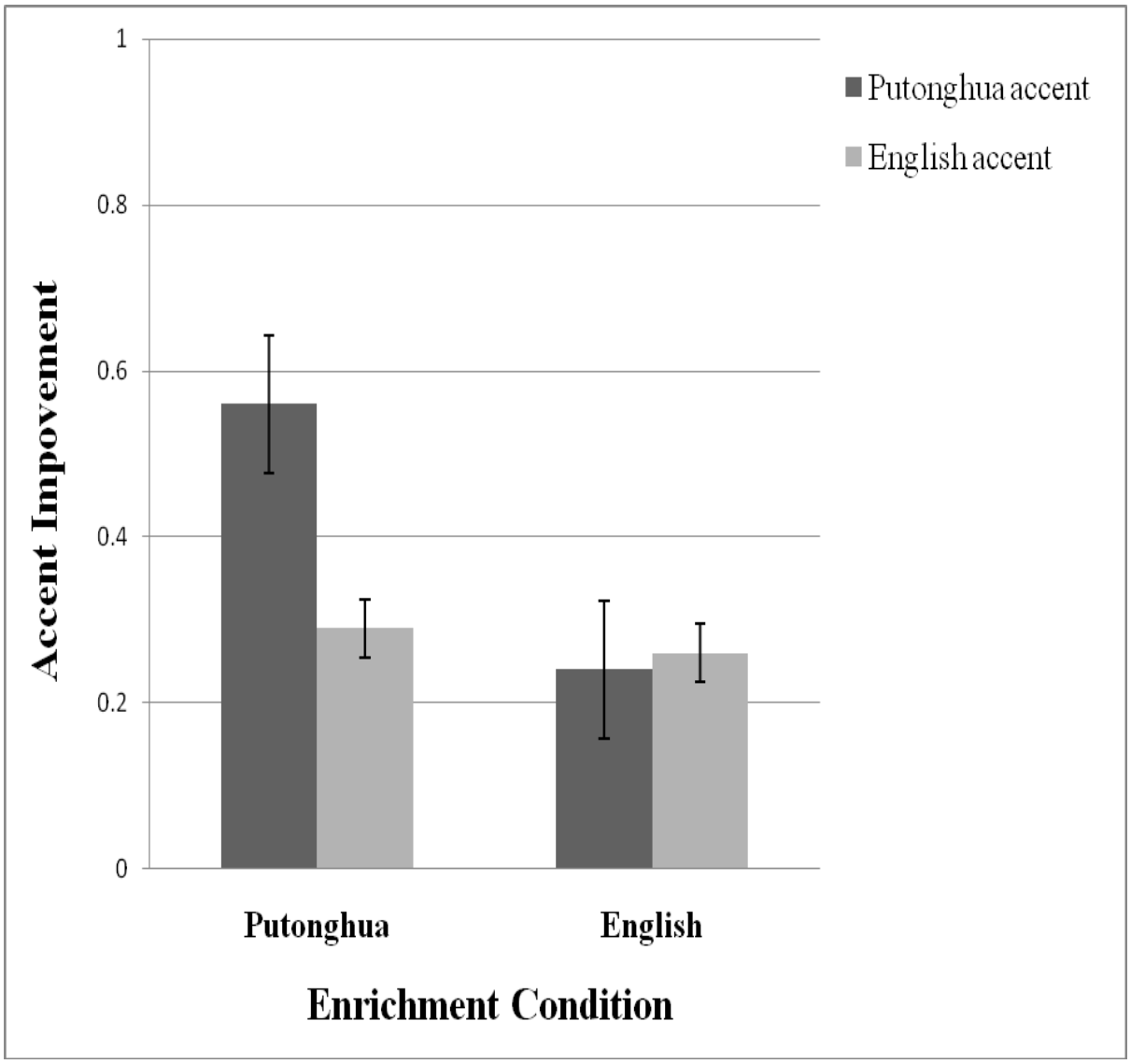
Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.

Oh, J. S., Au, T. K. & Jun, S.-A. (2010). Early childhood language memory in the speech perception of international adoptees. *Journal of Child Language, 37,* 1123-1132.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*, 261-283.

Pinker, S. (1994). *The language instinct*. New York: Morrow.
Ramsey, S. R. (1987). *The languages of China.* Princeton, NJ: Princeton University Press.
Rice, M. L. & Woodsmall, L. (1988). Lessons from television: Children's word learning when viewing. *Child Development, 59*, 420-429.
Rvachew, S. (1994). Speech perception training can facilitate sound production learning. *Journal of Speech & Hearing Research, 37*, 347-357.
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.
Slevc, L. R. & Miyake, A. (2006). Individual differences in second language proficiency: Does musical ability matter? *Psychological Science, 17*, 675-681.
Snow, C. E., 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, L. K. H., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language, 22*, 473-495.
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*, 579-590.
Todd, P. H. (1975). A case of structural interference across sensory modalities in second language learning. *Word, 27*, 102-18.
Todd, P. & Aitchison, J. (1980). Learning language the hard way. *First Language, 1*, 122-40.
Trainor, L. J., & Corrigall, K. A. (2010). Music acquisition and effects of musical experience. In M. Riess-Jones & R. R. Fay (Eds.), *Springer Handbook of Auditory Research: Music Perception* (pp. 89-128). Heidelberg: Springer.
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.
Werkers, 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*, 233-251.

Figure 1: Children's Mean Improvement Scores of Putonghua and English Accent in each Experimental Enrichment Condition



Appendix A: Song Titles for the Putonghua and English Enrichment Program

Putonghua Songs:

1. 嗡嗡嗡 wēng wēng wēng
2. 歡樂頌 huān lè sòng
3. 生日快樂歌 shēng rì kuài lè gē
4. 往事難忘 wǎng shì nán wàng
5. 當我們同在一起 dāng wǒ men tóng zài yì qǐ
6. 世界真細小 shì jiè zhēn xì xiǎo
7. 踏雪尋梅 tà xuě xún méi
8. 友誼萬歲 yǒu yí wàn suì
9. 小小姑娘 xiáo xiǎo gū niáng
10. 青春舞曲 qīng chūn wú qǔ
11. 跳舞歌 tiào wǔ gē
12. 媽媽好 mā mā hǎo
13. 我是隻小小鳥 wǒ shì zhī xiáo xiáo niǎo
14. 春神來了 chūn shén lái le
15. 倫敦橋 lún dūn qiáo
16. 小星星 xiǎo xīng xīng
17. 尋找愛 xún zhǎo ài
18. 恭喜恭喜 gōng xǐ gōng xǐ
19. 蠟燭 là zhú

English Songs:

1. More We Get Together
2. Down By the Bay
3. Brush Your Teeth
4. Robin the Rain
5. Five Little Frogs
6. I Wonder If I'm Growing
7. Aikendrum
8. Bumping Up and Down
9. Must Be Santa
10. Willoughby Wallaby Woo
11. Spider On the Floor
12. Baa Baa Black Sheep
13. Going to The Zoo
14. My Dreydel
15. Peanut Butter Sandwich
16. Five Little Pumpkins
17. Sharing Song
18. Mr. Sun
19. Old MacDonald Had a Band