

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/143200>

Please be advised that this information was generated on 2016-05-02 and may be subject to change.

Assessing birth language memory in young adoptees

© 2015, Wencui Zhou

Cover design: Wencui Zhou & Dirk van Kemenade

ISBN: 978-90-76203-68-3

Printed and bound by Ipskamp Drukkers b.v., Nijmegen

Assessing birth language memory in young adoptees

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. Th. L. M. Engelen,

volgens besluit van het college van decanen

in het openbaar te verdedigen op dinsdag 14 juli 2015

om 14:30 uur precies

door

Wencui Zhou (周文翠)

geboren op 20 oktober 1980

te Harbin, P. R. China

Promotor:

Prof. dr. Anne Cutler

Copromotor:

Dr. Mirjam Broersma

Manuscriptcommissie:

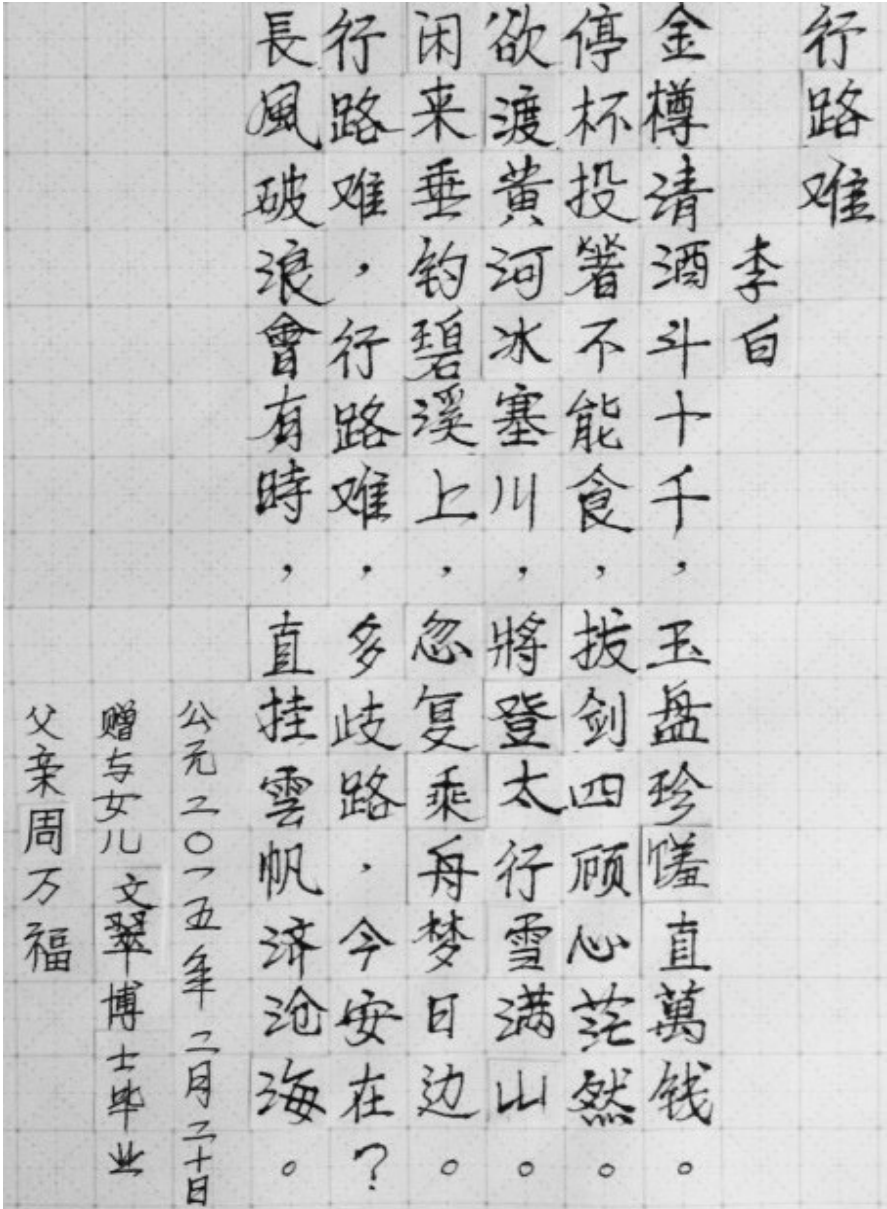
Prof. dr. Paula Fikkert (Radboud University)

Prof. dr. Femmie Juffer (Leiden University)

Dr. Christophe Pallier (INSERM Cognitive Neuroimaging Unit, CNRS)

The research reported in this thesis was supported by a grant from the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany, by the Donders Centre for Cognition, Radboud University, and by NOW Spinoza (“Native and Non-native Listening”).

Hard Road by Li Bai (Tang dynasty) hand written by my father Wanfu Zhou as a gift for my graduation.



Acknowledgements

As a traditional African proverb says, it takes a village to raise a child. To me, this PhD project is my precious baby. I would not have been able to fulfill this project without everyone's support over the years.

Among all those people, I would like to express my biggest appreciation to my two supervisors, Mirjam Broersma and Anne Cutler. Mirjam, I still remember our first meeting - the job interview - at the MPI. We had such a long and nice talk, definitely the best interview I have ever had. I have learnt so much from you over the years. You helped me fine-tune my research skills, and guided me through all the ups and downs. All the interactions we had in the past years shaped my academic career. Anne, you are like the head of a family in my eyes. Having you around, I always feel safe and sound at home. I know you will always be there for us, and yes you always have been. When our department was dismissed, I felt like a homeless child. However, your support has never stopped even from another continent. You have always motivated and encouraged me at critical moments. I cannot thank you enough for your kind and immediate support over the past years.

Next, I would like to thank my reading committee, Professor Paula Fikkert, Professor Femmie Juffer, and Dr. Christophe Pallier. Thank you all for reading my thesis and providing me with useful suggestions and comments, which helped improve the thesis tremendously.

Professor Monika Schmid, you opened the door to my academic career. I am grateful to have met you in Groningen. You introduced and guided me as I took my first steps in the field of research on international adoptees. Ever since then, I fell in love with academic research and have kept my mind active in this field for the past five years. Although we have had no more professional interaction since my Master's, your ongoing care supported me tremendously during my PhD. I truly appreciate that you 'checked on' my progress over the years, and recommended me academic opportunities in the past months.

Professor Carlos Gussenhoven, you are the best mentor in experimental phonology. I appreciated the opportunity to work with you on our 'perception of vowel duration' project. Your profound knowledge and superb research skills have always inspired me, and still continue to do so.

I would also like to thank Professor James McQueen (MPI & Radboud University), Dr. Makiko Sadakata (Radboud University), Professor Paula Fikkert (Radboud University), Associate Professor Aojun Chen (Utrecht University), Professor Holger Mitterer (University of Malta), Associate Professor Odette Schearenborg (Radboud University), Professor Natasha Warner (University of Arizona), Professor Aafke Hulk (University of Amsterdam), Professor Jesse Snedeker (Harvard University), and Associate Professor Paola Escudero (University of Western Sydney) for your kind suggestions and feedback on the experimental set-up of the project.

I am also grateful to Professor René Kager, Dr. Ao Chen, and Dr. Liquan Liu from Utrecht University, Associate Professor Emanuel Bylund (Stockholm University), Professor Fred Genesee (McGill University), and Associate Professor Nancy Eng (Hunter College CUNY) for your kindness to share your studies with me.

Many thanks to Professor Ping Li (Pennsylvania State University), Professor Yip Chi Wing Michael (The Hong Kong Institute of Education), and Professor Angel Wing Shan Chan (Polytechnic University of Hong Kong) for your suggestions and materials on the MacArthur-Bates Communicative Development Inventory (CDI) in Mandarin and Cantonese Chinese. Thanks also to Eva Wittenberg (Visiting fellow at Harvard University) who was so generous to share her graphical resources with me when I was preparing for my vocabulary study.

I am very thankful to the timely support from our technical department at the MPI, especially, Ronald Fischer who helped me program my complicated experiments, Reiner Dirksmeyer, Alex Dukers, Johan Weustink, and Rick van Viersen who provided the best technical support and arranged all my testing equipment, Ad Verbunt who explored Adobe Flash together with me, and Peter Nijland who converted thousands of AVI videos for my experiments. In the meantime, I also want to give my sincere appreciation to Tilman Harpe who created those lovely pandas and dinosaurs. Without your generous support, our participants would never have been able to play with those adorable pandas and dinosaurs.

I also want to thank our librarians, especially Karin Kastens and Meggie Uijen, who assisted me with EndNote, and ordered all the papers and books necessary for my research.

Thank you to the administration office, especially Paul Lommen and Rober van Thiel for arranging the institute car for us to collect data all over Holland, Marie-Luise Janssen-Ekat for handling all my traveling expense reimbursements, Angela Heuts for helping me renew my visa so many times, and Jan Achterberg for helping me move to several offices.

I would like to thank every 'family member' in the Language Comprehension department at the MPI, Anne, Mirjam, James, Odette, Eva, Holger, Falk, Frank, Andrea, Esther, Jiyoun, Sho, Katja, Patrick, Joost, Marijt, Matthias, Laurence and Neil (and others I might have forgotten to mention here). All the lively discussions during our meetings, and the happy times with 'interesting snacks' from different cultures reminded me what a great family we were/are. Rian, as our secretary, you have been extremely helpful to all of us over the years. I would also like to thank Antje Meyer who gave me a second home at the MPI, the Psychology of Language department. Thanks to Evelyn who arranged all my paper work. Special thanks to everyone at the Psychology of Language department, particularly Agnieszka, Alma, Suzanne, Svetlana, Florian, Maartje and Will. I must also thank the International Max Planck Research School (IMPRS) for Language Sciences, especially Els den Os, Rachel Sheer, and Dirkje van der Aa, for organizing all the interesting and useful events for our PhDs, and helping us enroll in useful courses at the University. Dirkje, thanks a lot for your kind support, particularly at the final stage of my PhD. Thank you for checking on my progress, keeping me up to date with job opportunities, and encouraging me at difficult moments.

My special thanks here go to my two gorgeous paranymphs, Zeshu Shao and Alina Lartseva. I am so happy to have you two standing by my side on this special day. You two are not only my good colleagues who have been tremendously helpful, but also my very good friends who have been enormously supportive. We shared laughter and tears over the years. Today, I am so pleased to share such an important moment in my life with you two.

Thanks to all my research assistants (Lotte de Haas, Neeltje Vaes, and Ilona Wissink), who traveled all over the Netherlands together with me testing over a 100 children. Thank you, Lies Cuijpers, Wendy van Ginkel, and Denise Moerel, for helping me segment the thousands of stimuli for my experiments.

Thank you all, my patient and supportive officemates, Matthias, Marijt, Joost, Jiyoun, Katja, Patrick, Laurence, and Sho. Over the past years, I moved to four different offices. I

was lucky to have shared the same office with you. Thank you all for the chit-chat, and for your help with my numerous work and/or life related questions.

During the preparation of my experiments, the members in the First Language Acquisition Group provided me useful feedback, in particular, I would like to thank Helen and Sho who shared their expertise in child language research, and gave me suggestions on how to make animated video games. When I was lost in R, Joost, Sho, Agnieszka, and Malte, offered me useful support. Thanks also go to Dan Dediu, James McQueen, Atsuko Takashima and every member in the Sound Learning Meeting group, who gave me useful suggestions to help me better understand the results of my experiments.

Millions of thanks go to my lovely friends in Nijmegen. Carmen, you are such a great friend. You have always been there for me whenever I need you for both work and friendship. We shared so much laughter and tears, which will be one of the most beautiful memories in my life. Salomi, you have been a fantastic friend. You have been super supportive to me both at work and in personal life. I appreciate your always objective and critical suggestions on my project. When I was isolated in China during my field trip, it was you who bridged the gap between me and the life in Nijmegen. At the final stage of our PhDs, when everyone was leaving, you were so kind to invite us to visit your home country. I was so touched by your family's hospitality. That was the best holiday I have ever had. Sho, you are such a wise and sweet person. I always admire the way you organize your work and life. Every time when I talked to you, you always motivated me. I am so grateful for the surprise birthday party you organized for me in the first year of my PhD. That was the best birthday I have ever had. Thank you so much for organizing the PhD dance. What a crazily happy day we had! Every time when I watch the video again, I cannot help laughing. What a lovely memory! I also enjoyed all those nights out for nice food and drinks at your place. We shared so many girl talks, life-stories, and sometimes hot debates on small interesting things, which I miss so much these days. Jiyou, you are not only my friend, but also my 'sister in arms'. We had been working on similar topics, so we could understand each other so well, particularly during the difficult data collection periods. It was so nice to have you around. I enjoyed our time together both at work and during the gatherings besides work, and of course during our holiday in Italy. Thank you so much for your sweet support and encouragement especially during my down moments. Jana, my friend and housemate, thank you for being considerate

and supportive all the time. I appreciate that you and Karen proofread my chapters and job application material. I miss the girl talk we had, the sweet cookies you made, and those home-made pizza nights we had. I loved your 'Christmas socks' full of candies, which motivated me the whole gloomy December. Rósa, I still remember how much fun we had when we watched Eurovision song contest at your place in 2010. I enjoyed all the nights out we had for cocktails or music events. Thank you very much for your friendship over the years. Katja, you were my officemate during the final period of my PhD. Although it was only a few months, I can never forget how much your cheerful personality brought laughter and joy to the office. When I was stressed, your 'do one thing at a time' helped me a lot. Loan, I miss the Vietnamese fresh spring rolls you made for us. It was such a great memory of our Italian holiday. Lin, you are a very sweet friend. I am so happy to have met you in Nijmegen in 2010, and have become good friends with you. Without you, my experiments in Beijing could never have been finished so quickly and smoothly. Every time when you visited Holland, you always brought me something from home (China). Huadong, a friend I have great faith in. Thank you for all your kind support in the past years. I am grateful to you and Lin when my studio was broken into; you two cooked so many delicious Chinese dishes to comfort me. I also enjoyed our trip to Turkey, where the four of us (including Xiaoqian and Yi Wan) had so much fun and so much nice Turkish food. Zeshu and Mingyuan, you two have been enormously supportive. I am grateful for your always kind suggestions both for work and for life. Thank you so much for hosting me at your place during Christmas and Chinese New Year in the past years. Thank you two for dropping by to accompany me when I was 'isolated' in my lonely office. Special thanks to little Peter, you brought me so much laughter and joy. Yan Gu, thank you for helping me find my first place to stay in Nijmegen. I miss all those wonderful dishes you made in our old house. Sophie, I enjoyed our cookie nights and IKEA trips so much. Thank you very much for helping me move with your cute white car. Meng, I always like our time together, including our shopping nights, running hours, Christmas cooking, and girl talk. I also appreciate those memorable gatherings and chats with Sarah, Christina, Antje, Helen, Steve, Nienke, Bohan, and Danchao, etc. Ewelina, Rebecca, Thordis, Saskia, Binyam and Beyza, because we were (all) at the final stage of our PhDs, it was great to cheer each other up, especially during the small chats we had in the corridor.

I also want to thank all my speakers who spent hours and hours in the small sound-proof booth in order to record enough stimuli for my experiments. In particular, thanks to Zeshu, Danchao, and Fengyuan, for their patience to record my stimuli so many times; thank

you Jie Wang, Jiaying Yu, Qingyuan Guo, and Summer Ruan for traveling for such a long distance several times to record my stimuli. Thanks to Mirjam, Odette, and Maartje for happily reading my stimuli in any style I requested.

Thanks also to all my friends outside of work, especially Maja Kutlaca, Jana Pejcinovska and Alok Chaubal, you provided me a new home whenever I visited Groningen. I miss the joyful times we spent together. John (Yannis Vasilakis), it is always nice to talk to you on Skype, updating each other on life, and motivating each other whenever we needed to. Xuejing Wu, Wei Wang, Xuan Liu, Shiping, Fiona, and Huisi, my friends in China, thank you so much for your considerate messages over the years. Even though I sometimes delayed replies for weeks, you still treat me so nicely. Even though we have had little chance to catch up, my friendship towards you guys has never been stronger.

I also owe many thanks to all Chinese and Dutch children, and their parents. Thank you for your time, and enthusiasm in participating in my long experiments, and for your kindness to help me recruit more children for my project. Without your participation, this project would not have been realized. Thanks also to all the Chinese and Dutch adult participants, who took part in my experiments with patience and passion.

With this opportunity, I would also like to express my appreciation and my love to my parents, my younger brother, my sister-in-law, and my seven-months-old niece. Thank you for your unconditional love and supports all the time. Your love and care have always motivated me to become a better person. 我爱你们!

Last but not least, I would like to give my special thanks to my beloved fiancé, Dirk. You cannot imagine how much your love and support mean to me all the time. Your enormous patience and care always remind me how lucky I am to have you in my life. I love you so much. I am so looking forward to our bright future together (with our little Dirk junior☺).

Table of contents

Acknowledgements..... V

Chapter 1: General introduction..... 1

 1.1 Linguistic memory 2

 1.2 Birth language retention 4

 1.3 Birth language relearning 7

 1.4 Research questions of the present study..... 11

 1.5 Chinese adoptees 12

 1.6 Cantonese and Mandarin Chinese 13

 1.7 Outline of the thesis..... 19

Chapter 2: Methodology 21

 2.1 Participants 21

 2.2 Materials..... 28

 2.3 Procedure..... 30

 2.4 Training: Method and Results 34

 2.5 Questionnaire on participants' characteristics 38

Chapter 3: Perception of the phonological contrasts of the birth language 41

 3.1 Experiment 1: Cantonese perception..... 41

 3.2 Experiment 2: Mandarin perception..... 50

 3.3 General discussion..... 56

Chapter 4: Production of the phonological contrasts of the birth language 63

 4.1 Experiment 1: Cantonese production 63

 4.2 Experiment 2: Mandarin production 77

 4.3 General discussion..... 88

Chapter 5: Residual memory of birth language vocabulary 93

 5.1 Experiment 1: Cantonese vocabulary..... 93

 5.2 Experiment 2: Mandarin vocabulary 100

 5.3 General discussion..... 103

Chapter 6: General discussion 107

 6.1 Long-term benefit of early linguistic experience 107

 6.2 Effects of AoA and Age (and LoR) 109

 6.3 Effects of Siblings and China Visits 110

 6.4 Impacts of monolingual versus bilingual input 111

 6.5 Conclusion..... 112

References..... 114

Appendix A..... 121

Appendix B..... 123

Appendix C..... 126

Appendix D..... 131

Appendix E..... 139

Appendix F..... 141

Nederlandse samenvatting..... 161

中文概述..... 165

Curriculum Vitae..... 168

Publications..... 169

MPI series in Psycholinguistics..... 170

Chapter 1: General introduction

Will international adoptees forget their birth language completely after adoption, as disuse of a language leads to loss (e.g., Activation Threshold Hypothesis, Paradis, 1993)? There is anecdotal evidence that comes from international adoptees, reporting that they do not remember their birth language anymore some time after adoption (see discussions on the adoptees' forum <http://www.fosteringfamilies.com/ref2.html>). Moreover, assumptions have been made (e.g., Glennen & Masters, 2002; Glennen, 2002) that birth language would undergo attrition or even complete loss when international adoptees start to learn their adoptive language, because exposure to the birth language is completely cut off. However, those anecdotal reports and assumptions have not been bolstered by empirical evidence. How much do we know so far about birth language development in international adoptees?

Previous research investigated whether all traces of birth language memory are lost, or whether those memories might just be dormant and refreshable with re-exposure. In those studies, the adoptees investigated were either adults or in their later childhood, several decades or at least ten years after adoption (e.g., Choi, 2014; Hyltenstam, Bylund, Abrahamsson, & Park, 2009; Oh, Au, & Jun, 2010; Pierce, Klein, Chen, Delcenserie, & Genesee, 2014; Singh, Liederman, Mierzejewski, & Barnes, 2011). No studies so far have addressed this issue in a much younger population: Adoptees who have been separated from their birth language for a relatively short period of time. Moreover, there is no research that has investigated the phonology (the aspect of a language that children develop the earliest) of the birth language in young adopted children. Finally, it is still unknown whether there is any conscious knowledge of birth language vocabulary. These are the questions the present project sets out to investigate. Thus, the current project is likely to be the first empirical study to systematically investigate the memory of birth language in young adopted children with an experimentally designed training paradigm, with a focus in the domain of phonology.

In this project, two groups of children who were adopted from two different Chinese language areas are investigated, namely Cantonese Chinese and Mandarin Chinese. They are tested, through perceptual training, on phonological perception (see Chapter 3), phonological production (see Chapter 4), and birth language vocabulary (see Chapter 5) of their birth language. The young adoptees' performances in those three different aspects of their birth language are compared to those of four groups of age-matched non-adopted control children,

namely two groups of native Dutch children, one group of native Cantonese children, and one group of native Mandarin children.

1.1 Linguistic memory

Several previous studies suggest that linguistic knowledge, once learned, can be well maintained in memory despite a long period of disuse. A notion put forward by Bahrck (1984, p. 114) is that once learned, knowledge is preserved in "permastore" and remains accessible even after a long time of disuse. In agreement with this notion, American learners of Spanish were still able to re-access, particularly, the semantic contents of the language, even after up to 50 years' disuse of the language, although this knowledge had undergone rapid attrition in the initial few years (Bahrck, 1984). Moreover, the "Savings paradigm" proposed by Nelson (1978) also suggests that knowledge, particularly words, will not be completely lost once learned. Two studies from de Bot and colleagues provided converging evidence for this account. Two German students of Dutch were reported to be more accurate in recognizing the Dutch words they learned 30 years ago compared to new words they were instructed to learn during the experiment (de Bot & Stoessel, 2000). American learners of German who were studying German for the 3rd semester and Dutch learners of French who had studied French for 4-6 years during secondary education at the time of testing were also found to be more successful in recognizing the words they had learned previously than the words they were instructed to learn during the investigation (de Bot, Martens, & Stoessel, 2004). Both the "permastore" notion and the "Savings paradigm" suggest long-lasting memory of early acquired linguistic knowledge, particularly in the domains of semantics, and vocabulary. However, no direct evidence regarding phonology has been provided so far. Thus, the question that arises is: can we expect a similar re-accessibility to the sounds of a language that people had learned and/or were exposed to early in life?

At least for heritage language learners who had been exposed to their childhood language for a short period of time during their early lives, the answer is positive. Several studies show that heritage language learners are better at relearning the sounds of their childhood language later in life, compared to novice learners who had no prior exposure to the language.

Tees & Werker (1984) studied adult heritage language learners who had early experience with Hindi during their first or two years of life. The heritage language learners

were reported to perceive Hindi retroflex-dental contrasts better than novice learners who had no prior experience with the language. Such advantage of childhood language experience was demonstrated several days up to even one year after the heritage language learners had taken their Hindi class. Childhood hearers of Korean in (Oh, Jun, Knightly, & Au, 2003), four months after relearning the Korean language as adults, were also reported to perceive the three-way Korean stop contrast (plain, aspirated, and tense) better than novice learners with no prior experience of Korean. Finally, a subset of participants in (Bowers, Mattys, & Gage, 2009) who had learned either Hindi ($N = 1$) or Zulu ($N = 2$) during their childhood also showed an advantage in the perception of their childhood language contrasts after practice. Importantly, those heritage learners of Hindi and Zulu were reported to have no residual knowledge of their childhood languages before practice; re-exposure through practice seemed to have refreshed their memory of the sounds of their childhood languages. The consistent findings in the above-mentioned studies suggest that linguistic knowledge acquired in early childhood remains intact in memory/storage, even after a long time of minimum or no continuous exposure, and relearning helps to refresh the access to the memory/storage. In consequence, those heritage language learners were able to learn their childhood language better than the novice learners.

Taken together, research on second and heritage language learners so far suggests that linguistic knowledge, such as semantics, vocabulary, and even phonology, acquired early in life can be accessed again despite a long period of disuse. However, note that both the second and heritage language learners in the previous studies had received continuous exposure to either their second or their childhood language although it was reported to be minimal (if any) later in life. There is a possibility that the continued exposure may have helped the second and the heritage language learners maintain their memory of their second or childhood language. Interestingly, ample evidence from studies in animals' visual and fear memories suggest that even under conditions of complete isolation from early on, such early memories are well retained, and can be retrieved with re-exposure (e.g., fear memory: Bouton & Peck, 1989; Bouton, 1984, 1993; Hardt, Wang, & Nader, 2009; visual memory: Hofer, Mrsic-Flogel, Bonhoeffer, & Hübener, 2006; Knudsen, 1998, 2004). Will international adoptees, completely cut off from their birth language after adoption, be able to remember their birth language immediately when they are exposed to the language again? How much do we know from empirical research so far?

1.2 Birth language retention

From the point of view of linguistic development, international adoptees are unique language acquirers. They acquire their birth language as first language learners do during their early childhood, but experience a sudden cut-off at adoption, and are faced with the task of learning a new language. So far a substantial number of studies have sought to understand the development of the adoptive language in international adoptees, particularly English language (e.g., Geren, Shafto, Snedeker, & Geren, 2009; Glennen, 2002, 2005; Krakow & Roberts, 2003; Karen Pollock, Price, & Fulmer, 2003; Roberts et al., 2005; Snedeker, Geren, & Martin, 2005; Snedeker, Geren, & Shafto, 2007). In general, adoptees were reported to gradually catch up with their non-adopted peers, although with an initial delay of one or two years' post-adoption (e.g., Glennen & Masters, 2002; Roberts et al., 2005). However, less is known with respect to the development of their birth language after adoption.

A neuroimaging study by Pallier and colleagues (2003) studied eight adult Korean adoptees who had been adopted into French-speaking families for an average of 26;8 years. In a speech segment detection task with Korean speech stimuli, the adoptees showed similar cortical activation patterns to those non-adopted French control participants. Moreover, similar brain regions were activated when they heard speech stimuli in all four different languages, including their birth language (Korean), their adoptive language (French), and two other foreign languages (Japanese and Polish). Two additional behavioral tasks, namely language identification and word recognition, also confirmed the adoptees' insensitivity to their birth language. It is striking that all eight adoptees were adopted after the age of three years, and three of them were even adopted at an advanced age (between seven and eight years old), but they still failed to show any sign of their birth language memory. This study suggests that the early established birth language representations were lost in the adoptees when no continuous input/exposure was available.

However, note that the previous neuroimaging study mainly focused on global memory of the birth language rather than memory of specific linguistic aspects. As the data of German Jews suggest, although the long period of disconnection of the birth language may affect the online processing of the language, the deeper linguistic repertoire may maintain undamaged (Schmid, 2002). This leaves open the possibility that adoptees might have retained knowledge of their birth language in some aspects or domains that most adoptees have acquired before adoption. Adoption generally happens during early childhood. Data

from the Centraal Bureau voor de Statistiek¹ shows that the majority of international adoptees in the Netherlands were adopted between the age of two and three years. At such a young age, which language aspect(s) are the adoptees expected to have developed the most?

Learning a language always starts with the sounds of the language. Evidence from research on first language acquisition shows that infants attune their perceptual abilities to the sounds in their first language with remarkable speed. According to the universal language timeline of speech perception development proposed by Kuhl (2004), infants attune their perceptual sensitivity to the vowels in their childhood language around six months old, and to the consonants around eleven months old. As for suprasegmentals, infants as young as four months old are sensitive to tones (Gao, Shi, & Li, 2010; Mattock, Molnar, Polka, & Burnham, 2008). By the time they reach their first birthday, infants have perceptually acquired their native phonological repertoire (Cutler, 2012).

This suggests that birth language phonology is the most promising aspect of birth language development for researchers to investigate in international adoptees. There is a possibility that adoptees have retained memory of their birth language sounds, since those are the first elements of birth language that they have heard, and probably the elements that they have best acquired before adoption. Is this indeed the case?

Findings of previous studies so far have been inconsistent. A recent fMRI study shows that early representations of the birth language can be well maintained in brain (Pierce et al., 2014). In that study, adopted Chinese children, on average 12.6 years after adoption, showed similar neural patterns to their Chinese-French bilingual peers when discriminating Mandarin tones. This finding contrasts to the previous neuroimaging study by Pallier et al. (2003) in which no trace of birth language memory was detected in the adult Korean adoptees' brains. Moreover, it is also different from the behavioral study by Ventureyra, Pallier, & Yoo (2004) in which eighteen adult Korean adoptees did not differ from their non-adopted French peers in their perception of a Korean three-way contrast of voiceless stops (namely, fortis, aspirated, and lenis) at the velar and bilabial places of articulation, and a Korean two-way contrast of denti-alveolar voiceless fricatives (namely, plain versus fortis). The difference of those results discussed above is striking given the fact that the adopted Chinese children in Pierce et al. (2014) were adopted at a much younger age, i.e., on average 12.8 months, compared to the adult Korean adoptees in Pallier et al. (2003) and Ventureyra et

¹ <http://www.cbs.nl/nl-NL/menu/themas/bevolking/publicaties/artikelen/archief/2011/2011-3449-wm.htm>

al. (2004) who were all adopted after the age of three years (several of them were even adopted at an advanced age, e.g., 9 years old). However, note that those adult Korean adoptees were tested at least two decades after adoption, it might be just too difficult for them to show any sign of birth language remnants after such long period of isolation from their birth language. So far evidence at the neural level suggests a well-maintained phonological memory of birth language, particularly in older adopted children; however, it is still unclear whether such evidence is also present at the behavioral level in younger adopted children?

To the best of our knowledge, only two case studies followed individuals soon after adoption. Instead of showing a well-preserved memory of the birth language, those case studies demonstrated how quickly the adopted children lost the access to their birth language. Nicoladis & Grabis (2002) studied a Cantonese girl (age of adoption: 17 months) four weeks after she was adopted by an English speaking family in Canada. They videotaped the Cantonese language interactions between the girl and the experimenters during five sessions over a period of 75 days. By the end of the investigation, the girl was reported to have lost her global communicative skills in both production and comprehension. Isurin (2000) studied vocabulary retention over a period of two years with a Russian girl who was adopted at the age of nine. At adoption, the girl's proficiency in Russian had been well developed. Except for the interactions with the Russian experimenter during the first three months (maximally an hour each month), the girl received no more exposure to Russian after adoption. Four months after adoption, the girl managed to successfully produce all Russian words in a Russian vocabulary test. Ten months after adoption, the girl could still reach 85% accuracy in the same vocabulary test. However, two years after adoption, the girl was only able to retrieve 40% of those Russian words. Not only production skills, but the girl's comprehension skills were also reported to decline tremendously over time. Both cases suggest a gradual loss of birth language after adoption; however, as they are case studies, no generalizations can be made. Most importantly, they assessed either global communicative skills of the birth language or the productive retrieval of birth language vocabulary. No study so far has looked at the phonological memory of birth language in young adoptees.

So far studies on birth language development in international adoptees provide contradictory findings. One shows well-maintained memory of the birth language, but the majority suggests a total or gradual loss of the birth language. A crucial question is whether it is simply too difficult to re-access to the knowledge of birth language due to non-continuous exposure and input. Note that in Pierce et al. (2014), participants were trained on the

perception of the Mandarin tones before being scanned. Although there was no difference between the adoptees and the French controls in a behavioral tone discrimination task, the perceptual training may have re-activated the adoptees' tonal memory of their birth language, thus demonstrated a native pattern of tonal processing at the neural level. An interesting question raised is: can re-exposure, such as relearning the language, help reactivate the “rusted” memory of birth language? As introduced in section 1.1, the positive evidence from those heritage language learners inspires us to consider whether relearning will also help international adoptees, who had experienced a complete cut-off from their birth language, re-access the memory of their birth language. How much do we know in this matter so far?

1.3 Birth language relearning

1.3.1 Perception of the phonological contrasts in the birth language

Several studies on adult and teenage adoptees suggest that relearning helps adoptees regain sensitivity to their birth language sounds. This is the case even though the adoptees had been isolated from their birth language for periods ranging from 10 years up to several decades. (Hyltenstam et al., 2009) assessed 21 adult Korean adoptees on their perception of two Korean contrasts, namely the three-way stop contrast at velar, bilabial, and labiodental places of articulation, and the plain versus fortis alveolar affricate contrast. Before the test, the adoptees had received formal Korean language training for an average of 2;1 years, and their non-adopted peers for an average of 4;1 years. The result showed that although one third of the adoptees scored higher than the best-performing control participant, there was no significant difference between the adoptees and the non-adopted controls. Most strikingly, the adoptees were outperformed by the control participants in a grammaticality judgment test. However, note that the controls received four years' formal training of Korean language, which is much longer than the adoptees. This may explain the significant performance of the control participants on the grammaticality judgment test. Nevertheless, this study was the very first attempt to investigate birth language retrieval through relearning. Because the two groups of participants were not well matched in terms of length of formal Korean language training, however, any advantage that the adoptees might have had was less likely to be detected.

Oh et al. (2010) continued with 12 adult Korean adoptees who had taken a Korean language class (five days a week for 50 minutes per day) at the University of Minnesota,

USA. During the second week of the Korean course, the adoptees were tested with a phoneme identification task on the Korean three-way stop contrast (at velar and denti-alveolar places of articulation). The results showed that although the adoptees were no better than the novice learners at identifying the phonological contrast in the overall analysis, they outperformed the controls in selective types of phonemes, such as lenis and aspirated stops. In contrast to the adoptees in previous investigations (e.g., Pallier et al., 2003; Ventureyra et al., 2004), this group of 12 adoptees were adopted at a younger age (Mean = 2;4 years), and had also experienced a long break (around 20 years) from their birth language after adoption. A short training of the birth language was nonetheless sufficient to revive part of the preserved phonological memory of the birth language. This positive finding seems to suggest that some phonological knowledge of the birth language may have been maintained even after long-term lack of use, and relearning can help re-access the deeply buried memory of the knowledge. This finding has been replicated in a recently completed PhD project by Choi (2014). Twenty-nine adult Korean adoptees were trained and tested on the perception of the Korean three-way stop contrast (at velar, denti-alveolar, and bilabial places of articulation), and showed a robust advantage in perceiving the contrast after a few training blocks ($N = 300$ trials). Another study on a group of teenage Indian adoptees who had been deprived of their birth language for an average of 10;3 years confirmed the positive effect of birth language relearning, particularly in the domain of phonological perception (cf., Singh et al., 2011). Eight teenage adoptees from India were tested and trained on the perception of voiced and voiceless retroflex versus dental stop contrast in their birth languages. Although the adoptees failed to show any difference from their non-adopted peers before training, after one brief training session ($N = 64$ trials) they outperformed their non-adopted peers.

In sum, studies on birth language relearning so far suggest that relearning even when limited to a short experimental training session is able to refresh the adoptees' phonological memory of their birth language, although they may have forgotten the language several years or decades after adoption. Birth language disuse may lead to forgetting; however, it does not entail a complete loss. Thus, can we expect a similar robust advantage of birth language relearning in a much younger population of international adoptees? That is what the Chapter 3 investigates.

1.3.2 Production of the phonological contrasts in the birth language

As put forward in the previous section, relearning the birth language helps adoptees perceive the contrasts of the language. It makes us wonder whether such advantage is evident in the

domain of production as well. That is, will the international adoptees be able to produce their birth language contrasts well after relearning?

To the best of our knowledge, only one PhD project (Choi, 2014) looked into this issue systematically. Twenty-nine adult Korean adoptees were tested on the production of the Korean three-way stop contrast (at velar, denti-alveolar, and bilabial places of articulation) before and after perceptual training. Results showed that the perceptual training helped the adult adoptees pronounce their birth language phonemes in a native-like manner. Strikingly, even though some adoptees were adopted even before they were able to speak (around 3 months old), the adopted sample as a whole demonstrated an overall privilege in the phonological production of their birth language.

Note that very little evidence on the advantage of phonological production of the birth language in international adoptees is available so far; however, research on the production of the childhood language in heritage language learners may lend us some knowledge to better understand this matter.

Several studies from Knightly, Au, and colleagues investigated heritage language learners of Spanish. “Heritage learners” have overheard and/or spoken a family language during early childhood. These studies consistently reported an advantage of speaking the childhood language with a reliable accent in those childhood overhearers and/or speakers when they relearn the language later in life. Knightly, Jun, Oh, & Au (2003) studied 15 childhood overhears of Spanish who were reported to hear Spanish regularly but speak Spanish minimally (limited to short phrases and words) between birth and the age of six years. After the age of six, they received much less exposure to Spanish. At the time of testing, they were on the second year of a Spanish language course at UCLA, and had re-learned Spanish for four or five years. The results showed that the childhood overhearers spoke better Spanish in terms of voice onset times (VOTs), degrees of lenition, and accents, compared to novice learners. A subset of that sample, i.e., 11 childhood overhears of Spanish who had been previously studied by (Au, Knightly, Jun, & Oh, 2002) also confirmed the production advantage in both VOTs and accents for the childhood overhears.

However, unlike the childhood overhears of Spanish, six childhood overhearers of Korean in Oh et al. (2003), four months after relearning Korean, failed to utter the Korean three-way stop contrast at the denti-alveolar place of articulation better than the novice learners, although they had heard Korean speech regularly before the age of five ($M = 40.3$ hours per

week). In contrast, the 15 childhood speakers of Korean in the same study who had spoken Korean regularly for at least three years before the age of seven ($M = 28.6$ hours per week) outperformed the novice learners in both VOTs and accents. As Oh and colleagues explained, two possible reasons may have contributed to the discrepancy in the results between the childhood overhears of Spanish and Korean. First, the childhood overhearers of Spanish relearned their childhood language for a much longer time compared to the childhood overhearers of Korean. Second, both heritage language learner groups were recruited from an American English speaking environment. The Spanish two-way stop contrast is more similar to the English two-way stop contrast, and hence is likely to be easier to learn, compared to the Korean three-way stop contrast.

A study from Au, Oh, Knightly, Jun, & Romo (2008) on childhood overhearers and speakers of Spanish seems to support this conclusion. 20 childhood overhearers who had overheard Spanish regularly before the age of six ($M = 6.2$ hours per week) (13 of them participated in Au et al., 2002 and; Knightly et al., 2003), and 10 childhood speakers who had spoken Spanish regularly ($M = 30.8$ hours per week) for at least three years before the age of seven were tested. Both groups of learners received significantly less exposure to Spanish after their childhood until they started relearning the language at the age of 14. At the time of testing, they were all on the second year of a Spanish language course at UCLA. Results from the accent rating task on the Spanish 2-way stop contrast, and evaluations of their natural speech, consistently showed that the pronunciations of both childhood overhearers and speakers were more highly rated by native listeners than those of the typical late L2-learners.

Early hearing and/or speaking experience of Spanish and Korean provides long lasting benefit and facilitates speaking the language later in life. This positive finding may leave an impression that a similar advantage in phonological production can be expected in international adoptees. However, note that the heritage language learners in the above-mentioned studies had received continued exposure to their childhood language later in life, though to a limited or minimal extent (e.g., a couple of hours a week). There is a possibility that the continuous exposure may have helped childhood overhearers and/or speakers maintain the phonological sensitivities to their childhood language, and thus assisted them to speak the language when relearning it. In contrast, international adoptees are normally adopted at an early age, most of them may have not yet started speaking before adoption, and they are completely cut off from their birth language since adoption. With such a disadvantage, compared to the heritage language learners, will the international adoptees be

able to produce their birth language contrasts better than their non-adopted peer before, and/or after relearning? That is what the Chapter 4 studies.

1.3.3 Residual memory of birth language vocabulary

Parallel to the sounds of the birth language, vocabulary is another aspect that international adoptees are widely exposed to before adoption. Since the birth language has lost its communicative function after adoption, what may happen to the birth language vocabulary? Can international adoptees still recognize words in their birth language after adoption?

The German Jews showed well-maintained lexicon range and breadth of their L1 despite of several decades' non-contact with German (Schmid, 2002). However, some studies with international adoptees expected no conscious knowledge of the birth language vocabulary in the adoptees, particularly several decades after adoption. They tested birth language vocabulary mainly as a baseline indication of non-conscious knowledge of the birth language (e.g., Choi, 2014). Further studies explored the re-activation of childhood slang terms after relearning, but failed to show any remnants of memory of birth language vocabulary, probably due to insufficient re-exposure (e.g., Oh et al., 2010). One explicit study of the attrition/loss of the birth language vocabulary in an adopted child soon after adoption (Isurin, 2000) mainly concerned production. It might be the case that a residual memory of childhood words is still maintained at the perception level, but cannot be detected at the production level anymore years after adoption. Nevertheless, no systematic study so far has investigated the possible residual memory of birth language vocabulary in a large sample of adoptees. That is what Chapter 5 examines.

1.4 Research questions of the present study

The present study aims to fill the empirical gaps identified in the previous sections. Three research questions were investigated. First, are adopted Chinese children better at perceiving the phonological contrasts in their birth language, compared to non-adopted Dutch control children, before and after perceptual training (see Chapter 3)? Second, are the adopted Chinese children better at producing phonological contrasts in their birth language, compared to the non-adopted Dutch control children (see Chapter 4)? Third, are the adopted Chinese children still able to recognize birth language vocabulary several months/years after adoption (see Chapter 5)? To this end, adopted Chinese children in the Netherlands between the age of four and ten years were tested, and their performances were compared to age-matched non-

adopted Dutch control children, and native Chinese children. Importantly, for the purpose of replication, two groups of Chinese adoptees from two different Chinese language areas, namely Cantonese and Mandarin Chinese, were investigated. In total, six groups of children (one group of Cantonese adoptees, one group of non-adopted Dutch controls, one group of native Cantonese children, one group of mandarin adoptees, another group of non-adopted Dutch controls, and one group of native Mandarin children) were studied in the current study. Details of the participants are presented in Chapter 2.

1.5 Chinese adoptees

Why do we test Chinese adoptees in the present project? There are two main reasons. First, adoption from China to the Netherlands is still ongoing; testing young adopted children from China is feasible. Second, there is a large number of Chinese adoptees in the Netherlands. Up to 2011, more than 6000 Chinese children have been adopted by Dutch speaking families in the Netherlands (Centraal Bureau voor de Statistiek, 2012²). The first group of Chinese adoptees entered the Netherlands in 1991 (Meiling Stichting voor Adoptie en Projecthulp annual report 2010). Since then, China has become one of the major countries where international adoptees are adopted from in the Netherlands. According to the Centraal Bureau voor de Statistiek³, although international adoption to the Netherlands between 1995 and 2011 has undergone a rise-and-fall pattern, most number of adoptees in the Netherlands during that period have come from China. There is no statistical data on the exact number of Cantonese and Mandarin adoptees living in the Netherlands, however, the Cantonese and Mandarin regions seem to be the two major language areas from where most Chinese orphans have come. For instance, the annual reports between 2004 and 2012 provided by Meiling Stichting voor Adoptie en Projecthulp⁴ (one of the major adoption organizations in the Netherlands) show that more than 600 Chinese orphans were adopted through the organization during those nine years. Amongst these, nearly 18% came from the Cantonese Chinese area, and more than 36% were from the Mandarin Chinese area. The remaining 46% came from the other five Chinese dialects areas (i.e., Wu, Min, Hakka, Xiang and Gan).

Why are there so many Chinese adoptees? The main reason for the large number of Chinese adoptees is the one-child policy issued by the Chinese government in 1979 (Rocha

² <http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLNL&PA=80399NED&D1=0&D2=1&D3=0&D4=a&HD=131018-1215&HDR=T,G2&STB=G1,G3>

³ <http://www.cbs.nl/nl-NL/menu/themas/bevolking/publicaties/artikelen/archief/2011/2011-3449-wm.htm>

⁴ <http://www.meiling.nl/downloads/jaarverslagen>

da Silva, 2006). This policy was introduced to reduce the pressure from overpopulation and poverty. In the traditional Chinese culture, boys carry the family blood and name, and this led some Chinese parents to abandon first-born girls in order to have a chance to give birth to a boy at a subsequent attempt. As a result many girls were sent to orphanages. Another important reason is the undesirable economic situation in most Chinese families. For instance, Guangxi province where almost half of the region speaks Cantonese Chinese is one of the least economically developed regions in China⁵. Moreover, the financial situation in most Mandarin Chinese regions such as Inner Mongolia, Heilongjiang, Jilin, Liaoning, Henan, Anhui, Gansu, and Ningxia are not favorable either. As a consequence, many children, particularly girls, were abandoned while boys were kept in the hope that they would be strong enough to take up jobs to support the family.

1.6 Cantonese and Mandarin Chinese

Cantonese and Mandarin Chinese are two of the major groups of Chinese dialects used in P. R. China, together with Wu, Min, Hakka, Xiang and Gan (Norman, 2003). Cantonese Chinese was documented to be used in Hong Kong, Guangdong province, and southeast parts of Guangxi province (e.g., Nanning, Pingnan, Wuzhou, and Yulin, and etc.) (Bauer & Benedict, 1997; Hou, 2002). Note that besides Cantonese Chinese, Mandarin Chinese is also widely used in those regions since it is the official language in China. Thus it makes the Cantonese areas in China Cantonese-Mandarin bilingual areas⁶.

Mandarin Chinese, as the official language used in China, has been popularized across the whole country since 1950s, particularly since an edict enforcing its use was pronounced by the General Office of the State Council of China in 1956. According to the Language Atlas of China (Wurm et al., 1987) and Hou (2002, pp. 174–175), Mandarin Chinese itself has another eight regional varieties, namely Beijing Mandarin, Northeast Mandarin, Jilu Mandarin, Jiaoliao Mandarin, Zhongyuan Mandarin, Lanyin Mandarin, Southwest Mandarin, and Jianghuai Mandarin. In order to eliminate the complexity in the participants' data, in the current study, we only recruited Mandarin adoptees from the north and northeast of China where the Mandarin language has been rated to be more standard compared to the rest. Those regions include Beijing, Inner Mongolia, Heilongjiang, Jilin, Liaoning, Tianjin, Hebei, Henan, Anhui, Shaanxi, Shanxi, Gansu and Ningxia.

⁵ http://en.wikipedia.org/wiki/Guangxi_province

⁶ Note that there are also other dialects spoken in those regions, such as Hakka, together with Cantonese and Mandarin.

1.6.1 Chinese Materials

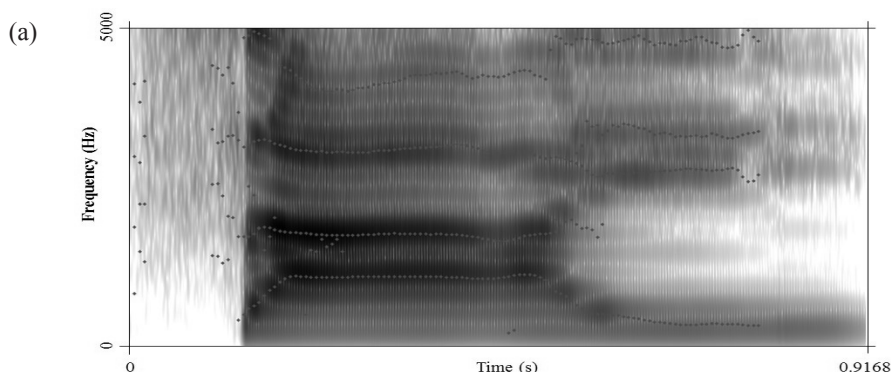
1.6.1.1 Phonological contrasts

For both the perception and production investigations, the Chinese affricate and tone contrasts were used. Both types of contrasts do not exist in the Dutch language, and were expected to be difficult for Dutch participants (as formalized in the predictions of the Perceptual Assimilation Model by Best & Tyler (2007); see examples of Dutch speakers perceiving English /æ/-/ɛ/ from Broersma (2005), and English speakers discriminating Mandarin alveo-palatal affricate vs. fricative from Tsao, Liu, & Khul (2006)).

1.6.1.1.1 Cantonese contrasts

1.6.1.1.1.1 Cantonese affricate contrast

There is one affricate contrast in Cantonese phonology, namely alveolar⁷ unaspirated affricate [ts] versus aspirated affricate [ts^h] (Bauer, 2010; Meng, Zee, & Lee, 2007; Zee, 1999). These are described as the voiceless stop [t] followed by the alveolar fricative [s], distinguished in aspiration (see Figure 1) (Bauer & Benedict, 1997). In most phonetic contexts the affricates are unpalatalized; when they are followed by the high front vowel [i:], and the front and central rounded vowels, [y:], [œ:] and [ø], they are palatalized to be [tʃ] and [tʃ^h] respectively (Bauer & Benedict, 1997). The palatalized affricates are very similar to the Mandarin alveopalatal affricates [tɕ] and [tɕ^h] (Bauer & Benedict, 1997), thus some Chinese linguists also annotate them as [tɕ] and [tɕ^h], particularly in Guangzhou Cantonese (Zhan et al., 1990).



⁷ There is no agreed terminology for the place of articulation for the Cantonese affricate contrast. Bauer & Benedict (1997) described it as dental affricate contrast, while Zee (1999) and Meng, Zee, & Lee (2007) referred it as alveolar affricate contrast. Bauer (2010) later used dental/alveolar affricate contrast. The present project sticks to the terminology used in Zee (1999) and Meng, Zee, & Lee (2007), namely alveolar affricate contrast.

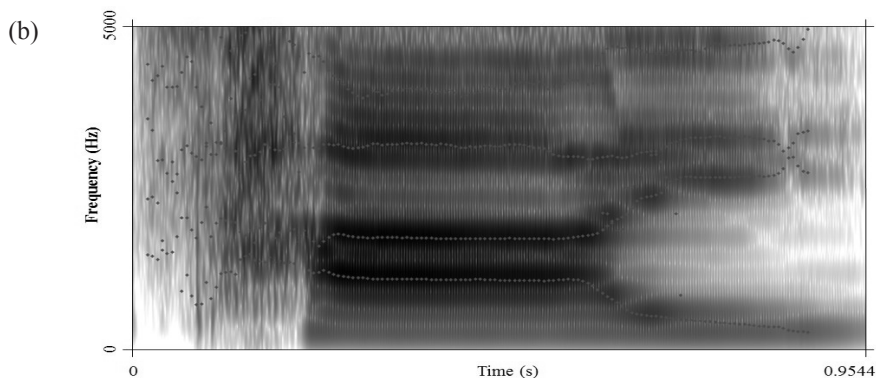


Figure 1. (a) Alveolar unaspirated affricate in the syllable [tsai], and (b) alveolar aspirated affricate in the syllable [tsʰai] produced by an adult female Guangzhou Cantonese speaker.

To the best of our knowledge, no research so far has documented the age of successful perception of the Cantonese affricates in native infants and toddlers. In production, they were reported to be the last phonemes acquired by native Cantonese children (So & Dodd, 1995). Some individuals were reported to be able to use the two affricates around the age of two (Tse, 1991; Tse, 1982), while bigger samples of Cantonese toddlers were found to successfully pronounce the two affricates only until the age of 4 or 5 years old (Cheung, 1995; Law & So, 2006; So & Dodd, 1995). Despite the late acquisition, affricates are salient in Cantonese phonology. The Cantonese version of early vocabulary inventory (translated from the Mandarin MacArthur-Bates Communicative Development Inventory by Hao, Shu, Xing & Li (2008) suggests that quite a few childhood words consist of affricates, which are reported to be perceptually acquired by more than 50% of the native children by the age of 12 months. For example, [tsøy] "嘴 mouth", [tsʰɔ̃] "坐 sit", [tsʰət min] "出面 outside", [tsʰɔ̃ŋ] "床 bed", and [hœŋ tsiu] "香蕉 banana", etc. Therefore, we expect Cantonese infants and toddlers to be sensitive to the Cantonese affricates in perception at least.

1.6.1.1.1.2 Cantonese tone contrast

There are six distinguishing lexical tones, particularly in open syllables, in Cantonese Chinese⁸ (Chen, Li, Shen, & Fu, 2001). They are Tone 1 (High-Level), Tone 2 (High-Rising), Tone 3 (Mid-Level), Tone 4 (Low-Falling), Tone 5 (Low-Rising), and Tone 6 (Low-Level), as displayed in Figure 2. The tone contrast tested in the current project is the rising tone

⁸ There are three additional lexical tones that only occur in closed syllables, ending with plosives [p], [t], or [k]. They are Tone 7 (High-Level), Tone 8 (Mid-Level), and Tone 9 (Low-Level). These three tones are similar to Tones 1, 3, and 6, respectively, in pitch and only differ in vowel length; According to the Romanization Scheme of the Linguistic Society of Hong Kong, the nine tones are normally combined into six tones (Chen et al., 2001).

contrast between Tone 2 (High-Rising) and Tone 5 (Low-Rising). These two tones are distinguished from each other on the height of their rises, as depicted in Figure 2. Tone 2 rises higher than Tone 5.

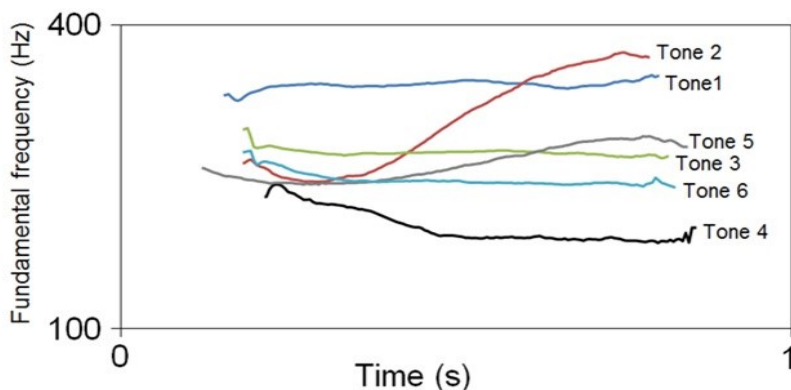


Figure 2. The six Cantonese tone contours in the syllable [si] produced by an adult female Cantonese speaker in isolation. [si-tone 1] means "poem", [si-tone 2] means "history", [si-tone 3] means "attempt", [si-tone 4] means "time", [si-tone 5] means "city", and [si-tone 6], means "trained person". Figure edited from Cutler (2012) with permission.

Very little is known about how early (or late) native Cantonese infants are able to discriminate the two Cantonese rising tones. In production, all six tones were reported to be acquired by the age of two, with Tone 2 earlier than Tone 5 (So & Dodd, 1995; Tse, 1978). The age of acquisition of Tone 2 is reported to be before the age of 1;8 years (1;3 years old in Tse (1991), 1;6 years old in So & Dodd (1995), and between 1;5 and 1;8 years old in Tse (1978)), while the age of acquisition of Tone 5 after the age of 1;8 years (1;8 years old in Tse (1991), 1;9 years old in Tse (1978), and between 1;8 and 1;11 years old in So & Dodd (1995)). The difference in the age of acquisition of the two rising tones can be caused by two reasons. Firstly, frequency difference: Tone 2 is one of the most frequent tones among the six tones, while Tone 5 is the least frequent tone (Leung, Law, & Fung, 2004). Secondly, contour similarity: Both tones are rising tones; their only difference is the height of rise at the end of the contour. As a consequence, a common confusion between these two tones has been reported in children between four and ten years old (Ciocca & Lui, 2003). The confusion is even noticeable in native adults (Ching, 1984; Lee, Kochanski, Shih, & Li, 2002). Nevertheless, lexical tones are integrated with all syllables in Cantonese phonology. Compared to non-native listeners, we expect Cantonese infants and toddlers to be more familiar with lexical tones since they have been exposed to tones since birth, or even before birth (given that auditory learning starts before birth, see Cutler (2012, pp. 259–301)).

1.6.1.1.2 Mandarin contrasts

1.6.1.1.2.1 Mandarin affricate contrast

Mandarin Chinese has three pairs of affricate contrasts: a retroflex affricate contrast: [tʂ] versus [tʂʰ], a dental affricate contrast: [tʃ] versus [tʃʰ], and an alveolo-palatal affricate contrast [tɕ] versus [tɕʰ]. Like the Cantonese affricate contrast, the three types of Mandarin affricate contrasts also contrast in aspiration. The present study focuses on the retroflex affricate contrast. With respect to pronunciation, the retroflex affricates [tʂ] and [tʂʰ] sound rather like the post-alveolar affricates [dʒ] and [tʃ] in English, respectively. Figure 3 displays spectrograms of the two affricates.

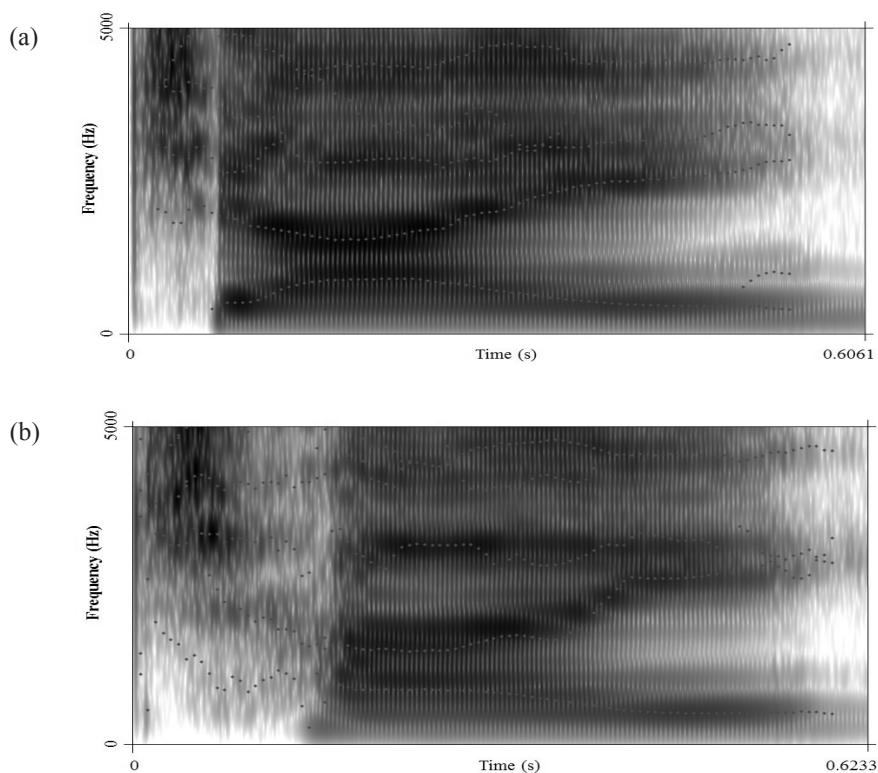


Figure 3. (a) Retroflex unaspirated affricate in the syllable [tʂai], and (b) retroflex aspirated affricate in the syllable [tʂʰai] produced by an adult female Mandarin speaker.

The same as the Cantonese affricates, no perception study has recorded the age of acquisition of the Mandarin affricate contrasts. However, in contrast to the Cantonese affricates, the Mandarin affricates were reported to emerge in native Mandarin toddlers' speech quite early (Zhu & Dodd, 2000). Zhu & Dodd (2000) reported that 90% of the native

Mandarin children started to use the retroflex affricates around the age of 2;1-2;6 years. By the age of 4;6 years, they were able to speak all the 21 Mandarin consonants at initial positions. This early acquisition of Mandarin affricates is likely to be attributed to their saliency in the Mandarin phonology, as demonstrated in the Mandarin children's vocabulary. Several childhood words in the Mandarin MacArthur-Bates Communicative Development Inventory which consist of retroflex affricates are reported to be perceptually acquired by 50% of native children at the age of one year (Hao et al., 2008). For instance, [tʂʰi] "吃eat", [tʂʰ uan] "床bed", and [tʂou] "粥porridge", etc. Thus, we expect Mandarin infants and toddlers to attend to retroflex affricates when they hear their native sounds.

1.6.1.1.2.2 Mandarin tone contrast

Four distinguishing lexical tones are used in Mandarin Chinese, namely Tone 1 (High-Level), Tone 2 (High-Rising), Tone 3 (Low-Dipping), and Tone 4 (High-Falling) (see Figure 4). The tone contrast investigated in the present project is between Tone 2 (High-Rising) and Tone 3 (Low-Dipping). This contrast differs from each other in pitch contours. Tone 2 is a smooth rise, while Tone 3 consists of two movements, falling and rising.

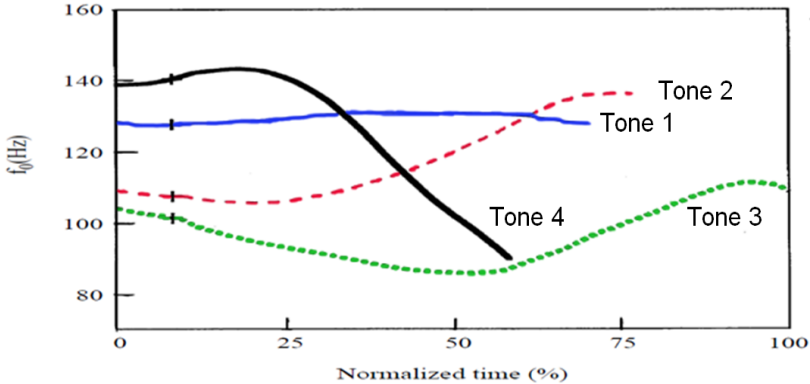


Figure 4: The four Mandarin lexical tones in the monosyllable “ma” produced in isolation. [ma-tone 1] means "mother", [ma-tone 2] means "hemp", [ma-tone 3] means "horse", and [ma-tone 4] means "to scold". Figure edited from Xu (1997, p. 7) with permission.

Evidence shows that Mandarin infants from four months of age were able to discriminate Mandarin Tone 2 and Tone 3 successfully (Gao et al., 2010). With respect to production, Zhu (2002) reported that children by the age of 1;10 years have acquired all four tones. In terms of the order and the age of emergence, Tone 2 occurs earlier (i.e., 1;3-1;4 years) than Tone 3 (i.e., 1;4-1;7 years) (Zhu, 2002). However, in terms of the age of

stabilization, both tones are reported to be mastered around the same age, i.e., 1;4-1;10 years (Zhu, 2002). Due to the similarity in the contours of the two tones, namely a rise in the end of each contour, confusion of the two tones is commonly observed in the early stage of acquisition for native children. Studies show that as early as one year old (i.e., 10-12 months); Mandarin children start confusing Tone 2 with Tone 3 (Tsao, 2008). This confusion persists throughout the first three and half years of children's childhood (Liu, 2007; Wong, Schwartz, & Jenkins, 2005). A study on Dutch infants also confirmed such confusion even with infants as young as six months (Chen & Kager, 2012). However, unlike Dutch infants, Mandarin infants and toddlers have sufficient input in their ambient phonology. Therefore, they are expected to be able to distinguish these two tones with less difficulty.

1.6.1.2 Chinese vocabulary

Based on the Mandarin MacArthur-Bates Communicative Development Inventory (henceforth M-CDI) provided in Hao, Shu, Xing & Li (2008), Chinese vocabulary were selected that was reported to be understandable by 50% of native Mandarin children at the age of 12 months. The selected vocabulary list consists of both verbs and nouns. Because we were unable to obtain a copy of the Cantonese MacArthur-Bates Communicative Development Inventory, the Cantonese vocabulary list was a translation of the Mandarin version by three native Cantonese speakers (*age* = 30, 23, 26 years) from Guangzhou, China. The word list with English translations is available in Appendix A.

1.7 Outline of the thesis

This dissertation consists of six chapters. Following this introductory chapter which has introduced previous accounts on linguistic memory, reviewed all available research on birth language development in international adoptees, pointed out what we are still missing in the current literature, and elaborated the research questions of the present project, Chapter two presents the general methodology and procedure of the study. It describes all participants investigated in the project, the materials, the procedure, and the perceptual training program of the study. Additionally, it details the two questionnaires i.e., general questionnaire and the N-CDI used in the project. Chapters three, four, and five are three central experimental chapters. Each chapter consists of two experiments in two different Chinese languages, namely Cantonese and Mandarin. Chapter three investigates the perception of the phonological contrasts in the birth language by adopted Chinese children. In Chapter four the

production of the phonological contrasts in the birth language is investigated. Chapter five tests the residual memory of the birth language vocabulary. The last chapter, Chapter six summarizes the results in all three experimental chapters, and discusses the main findings of the dissertation.

Chapter 2: Methodology

2.1 Participants

2.1.1 Cantonese experiment

Participants were 22 Cantonese adopted children, 23 non-adopted Dutch control children, and 22 non-adopted Cantonese control children. All were volunteers. They received a small monetary reward (presented to their parents) for their participation. Additionally, the Cantonese adoptees and the Dutch controls also received a collection of children's books.

The Cantonese adoptees (15 girls and 7 boys) were adopted from the Cantonese language areas in China by Dutch speaking families in the Netherlands. They were adopted between the age of 9 months and 4;6 years (i.e., 4 years and 6 months), with a mean age of adoption of 2;2 years ($SD = 1;2$). At the time of testing, they were between the age of 4;4 years and 10;10 years old, with a mean age of 7;5 years ($SD = 1;10$), and had lived in the Netherlands for a period of 1;5 to 9;11 years, with a mean length of residence of 5;3 years ($SD = 2;6$). Three adoptees had a cleft palate, and received surgeries before and/or after adoption, and speech therapy in the Netherlands. At the time of testing, their speech was reported to be intelligible by their adoptive parents and speech therapists. Another two adoptees had vision only in one of the two eyes. All others were reported to have normal speech, and normal or corrected-to-normal vision. All adoptees were reported to have normal hearing and motor control. They were recruited with the help of Dutch adoption organizations, travel agencies specialized in 'root seeking' trips for adoptive families, and through informal networks of adoptive parents.

The Dutch control children (12 girls and 11 boys) were born and lived in the Netherlands in their Dutch speaking birth families. When tested, they were between 4;7 and 10;6 years old, with a mean age of 7;8 years ($SD = 1;10$). All of them were reported to have normal speech, hearing, and motor control, and normal or corrected-to-normal vision. They were recruited from the siblings ($N = 5$), relatives ($N = 1$), or friends ($N = 2$) of the Cantonese adoptees, and through informal networks.

No children in either the adoptee group or the Dutch control group had received any Chinese language training before participation (where applicable: after adoption). By the time

of testing, none of them were reported to be able to understand any Cantonese Chinese by their parents. Additionally, as some dialects in the Dutch province Limburg use lexical tones (Gussenhoven & Peters, 2008; Gussenhoven, 1999), care was taken that none of the participants (had) lived in Limburg or were exposed to a Limburg dialect at home⁹. All children in both groups were reported to speak Dutch fluently. One adoptee and one Dutch control child were reported to be Dutch-English simultaneous bilinguals, with Dutch being dominant. All others were being raised monolingually. Participants in the two groups were well matched with respect to 'Age', 'Gender', the presence of Cantonese adopted siblings (i.e., 'Siblings'), visits to the Cantonese part of China (i.e., 'China Visits'), music training outside school (i.e., 'Music'), and Dutch vocabulary (as measured by the Dutch MacArthur-Bates Communicative Development Inventory) (i.e., N-CDI 3, Zink & Lejaegere, 2002). There were no significant differences between the groups with respect to any of those six variables ($ps > .05$) (see Table 1). The two groups were significantly different in terms of the highest educational level of the parent(s) ('Parent Highest Education'; see Table 1), $p < .05$ (which will therefore be controlled for in all analyses in the following chapters.)

As Table 1 shows, nine of the adoptees and five of the controls had a Cantonese adopted sibling. Three of those adoptees and two of those controls had two Cantonese adopted siblings each. Further, seven of the adoptees had a sibling from another part of China (where no Cantonese is spoken). One control participant had a Cantonese cousin, and two control participants had a Cantonese adopted friend. As also shown in Table 1, ten of the adoptees and four of the control participants had visited the Cantonese language areas of China (for adoptees: after adoption), with two visits for one adoptee (for a total of 24 days), and one visit for all others (average, for adoptees: 13 days; for controls: 19 days). Additionally, two of the adoptees had visited other parts of China (where other Chinese dialects instead of Cantonese are spoken). Seven of the adoptees and nine of the controls attended music classes outside school (Table 1), on average 11.9 months for the adoptees, and 15.8 months for the controls. Furthermore, Table 1 indicates the parent(s)' highest level of education (representing only the level of one parent, namely the one with the highest level of education of the two parents).

For all children, parents filled in the Dutch vocabulary checklist 'Oudervragenlijst N-CDI 3' (hence: 'N-CDI 3'), as an indication of their Dutch development, and to establish the lower bound of their Dutch development, particularly for the adoptees. The N-CDI 3 is the

⁹ The mother of one participant originated from Limburg, but she stated that she never used the Limburg dialect at home and had been living outside the Limburg area for 25 years.

highest level CDI, and was designed for children between the age of 30 and 37 months. Because all children in the present study were older than 37 months, as expected, non-adopted participants were all near ceiling. All adoptees had been exposed to Dutch for at least one year, and (with one exception; see below) were all near ceiling (in line with Snedeker et al., 2007). As can be seen in Table 1, the two groups of children did not differ significantly from each other on the N-CDI 3 percentile scores. The N-CDI 3 checklist consists of four parts: 1) active vocabulary, 2) phrases, 3) length of sentences produced, and 4) questions on comprehension, semantics, and syntax. The raw scores of each part were converted to percentiles, based on the normative tables in the N-CDI technical manual (Zink & Lejaegere, 2002). For each child, the average percentile of the four parts was calculated. All children had an average score above the 50th percentile (i.e. scored higher than 50% of their peers), with one exception¹⁰. The participant who scored below the 50th percentile was tested with another checklist, N-CDI/*Woorden en Zinnen*, which was designed for children between 16 and 30 months old, where his average score was at the 78th percentile.

We aimed to match the two groups as well as possible on all the above variables. Note, however, that due to the limited availability of Dutch children with adopted Chinese siblings, or who had visited China, the two groups could not be optimally matched on the variables Siblings and China Visits (although the proportions in both groups do not differ significantly on statistical tests). Further, the highest level of education of the parent(s) could not be optimally matched (and this led to a small but statistically significant educational advantage for the Dutch controls' parents).

The Cantonese control children (13 girls and 9 boys) were born and lived in Guangzhou city (China) in their Cantonese speaking birth families. They were tested between 4;3 and 10;5 years of age, with an average age of 7;3 years ($SD = 1;11$). All children were Cantonese-Mandarin fluent bilinguals, with Cantonese being used dominantly within the household, except for one child ($age = 4;3$ years) who understood but was not able to speak Mandarin well at the time of testing. Four children received music training outside school, for an average of 26 months. All of them were well matched with the Cantonese adoptees and the Dutch controls in Age, Gender, and Music ($ps > .05$). They were all reported to have normal speech, hearing, and motor control, and normal or corrected-to-normal vision. They were recruited from a local educational center for children and through informal networks.

¹⁰ One adoptee had an average percentile score of 45th (active vocabulary: 25th percentile scores; phrases: 50th percentile scores; length of sentences produced: 40th percentile scores; questions on comprehension, semantics, and syntax: 65th percentile scores). He was adopted at the age of 1;1 years and had lived in the Netherlands for 3;3 years by the time he was tested.

Table 1
Summary of statistics of the seven control variables for the Cantonese adoptees and the Dutch controls. See main text for explanation of variables.

Control variables	Cantonese adoptees			Dutch controls			Group difference
	N	%	M(SD)	N	%	M(SD)	
Age	22		7;5 (1;11)	23		7;8 (1;11)	$t(43) = -.388, p > .05$
Gender							$\chi^2(1) = 1.201, p > .05$
	Girl	15	68.2%	12	52.2%		
	Boy	7	31.8%	11	47.8%		
Siblings							$\chi^2(1) = 1.928, p > .05$
	Yes	9	40.9%	5	21.7%		
	No	13	59.1%	18	78.3%		
China Visits							$\chi^2(1) = 4.132, p = .057$
	Yes	10	45.5%	4	17.4%		
	No	12	54.5%	19	82.6%		
Music							$\chi^2(1) = .262, p > .05$
	Yes	7	31.8%	9	39.1%		
	No	15	68.2%	14	60.9%		
Parent Highest Education¹¹							$\chi^2(2) = 6.248, p < .05$
	Intermediate vocational ('MBO')			2	8.7%		
	Higher vocational ('HBO')	12	54.5%	5	21.7%		
	University ('WO')	10	45.5%	16	69.6%		
N-CDI 3		22	89.58 (13.37)	23	93.11 (4.84)		$t(26.203) = -1.167, p > .05$

¹¹ There are seven levels of secondary and tertiary education in the Netherlands (from low to high): VMBO-P (LBO), VMBO-T (MAVO), MBO, HAVO, VWO, HBO, WO.

2.1.2 Mandarin experiment

Twenty-six Mandarin adoptees, 26 Dutch controls, and 18 Mandarin controls participated in the Mandarin experiment. They also received a small amount of money and a collection of children's books for their participation, like the participants in the Cantonese experiment did.

The Mandarin adoptees (15 girls and 11 boys) were adopted from the Mandarin speaking areas in the north of China. They were adopted within the age range of 0;10-5;8 years, with an average age of adoption of 2;4 years ($SD = 1;2$). At the time when they were tested, they were between 4;1 and 10;10 years old, with an average age of 7;4 years ($SD = 1;9$), and had lived in the Netherlands for between 0;3 and 9;10 years, with an average length of residence of 5;0 years ($SD = 2;3$). One adoptee had a cleft palate, received recovery surgeries before adoption, and speech therapy in the Netherlands. One adoptee had only one hand. The rest of the adoptees were reported to have normal speech and normal motor control. All adoptees were reported to have normal hearing, and normal or corrected-to-normal vision. They were recruited in the same way as the Cantonese adoptees

The Dutch control children (13 girls and 13 boys) were also born and raised in their Dutch speaking birth families, like those in the Cantonese experiment. They were tested between 4;2 and 10;8 years of age, with an average age of 7;3 years ($SD = 1;11$). All were stated to have normal speech, hearing, and motor control, and normal or corrected-to-normal vision. Again as in the Cantonese experiment, they were recruited from the siblings ($N = 6$), and relatives ($N = 3$) of the Mandarin adoptees, and through informal networks.

Neither the adoptees nor the controls had attended any Chinese language courses before test (where applicable: after adoption), or understood Mandarin Chinese when being tested (with one exception¹²). Further, none of the children in either group had lived in the province of Limburg or had been exposed to the dialects spoken in Limburg. Both groups of children were fluent in Dutch (with one exception¹³). One adoptee was reported to be Dutch-German simultaneous bilingual, with Dutch being predominant at home and school. All other children were being raised monolingually. All children in both groups were well matched with regard to the control variables; Age, Gender, Siblings, China Visits, Music, Parent Highest Education, and N-CDI 3 (see Table 2), in contrast to the Cantonese experiment.

¹² One adoptee who was adopted three months before test was reported to have sometimes spoken Mandarin Chinese in daily communication.

¹³ The same adoptee who was adopted three months before test was reported to be only able to speak a few Dutch isolated words and short phrases.

As presented in Table 2, four of the adoptees and four of the control children had a Mandarin adopted sibling. Another two Dutch controls had two Mandarin adopted siblings. In addition, six of the adoptees had a sibling adopted from other parts of China where different Chinese dialects were spoken. Three control children had a Mandarin adopted cousin. As Table 2 also shows, five adoptees and one Dutch control child had once visited the Mandarin Chinese areas in China (for adoptees: after adoption) (average, for adoptees: 15.4 days; for Dutch control: 21 days), and one adoptee and one control child had visited it twice for a total length of 35 days each. Additionally, one adoptee had visited Hong Kong (where Cantonese is dominantly used) two times for six days in total. Six adoptees and nine controls attended music training outside school, on average 19 months for the adoptees, and 27 months for the controls. Table 2 also presents the highest educational level of the parent(s).

As in the Cantonese experiment, parents of all children in the present experiment also filled in the Dutch vocabulary checklist 'N-CDI 3'. As shown in Table 2, there was no group difference between the adoptees and the controls in their percentile scores of the N-CDI 3. All children in both groups reached an average score higher than the 50th percentile, with two exceptions¹⁴. Those two children who fell below the 50th percentile scores were tested with the N-CDI/*Woorden en Zinnen*, where they obtained an average score of 27th percentile, and 65th percentile, respectively.

The Mandarin control children (7 girls and 11 boys) were born and resided in Beijing city (China) in their Mandarin speaking birth families. When tested, they were between 4;1 and 8;3 years old, with a mean age of 6;6 years ($SD = 1;4$). All control children were reported to be Mandarin monolinguals. Eight of them received music training, for an average of 18 months. They were all well matched with both the adoptees and the Dutch controls in Age, Gender, and Music ($ps > .05$). All of them were reported to have normal speech, hearing, and motor control, and normal or corrected-to-normal vision. They were recruited through personal networks.

¹⁴ One adoptee had an average score of 2nd percentile in the N-CDI 3 (active vocabulary: 1st percentile score; phrases: 1st percentile score; length of sentences produced: 1st percentile score; questions on comprehension, semantics, and syntax: 5th percentile score). He was adopted at the age of 5;8 years, and had lived in the Netherlands for 3 months by the time when he was tested. The other adoptee had an average score of 34th percentile in the N-CDI 3 (active vocabulary: 30th percentile score; phrases: 25th percentile score; length of sentences produced: 65th percentile score; questions on comprehension, semantics, and syntax: 15th percentile score). She was adopted at the age of 3;0 years, and had lived in the Netherlands for 13 months at the time when she was tested.

Table 2
 Summary of statistics of the seven control variables for the Mandarin adoptees and the Dutch controls. See main text for explanation of variables.

	Control variables			Mandarin adoptees			Dutch controls			Group difference
	<i>N</i>	%	<i>M(SD)</i>	<i>N</i>	%	<i>M(SD)</i>	<i>N</i>	%	<i>M(SD)</i>	<i>p</i> -value
Age	26		7;4 (1;9)	26		7;3 (1;11)				$t(50) = .107, p > .05$
Gender										$\chi^2(1) = .310, p > .05$
	<i>Girl</i>	15	57.7%	13	50%					
	<i>Boy</i>	11	42.3%	13	50%					
Siblings										$\chi^2(1) = .495, p > .05$
	<i>Yes</i>	4	15.4%	6	23.1%					
	<i>No</i>	22	84.6%	20	76.9%					
China Visits										$\chi^2(1) = 2.364, p > .05$
	<i>Yes</i>	6	23.1%	2	0.7%					
	<i>No</i>	20	76.9%	24	92.3%					
Music										$\chi^2(1) = .843, p > .05$
	<i>Yes</i>	6	23.1%	9	34.6%					
	<i>No</i>	20	76.9%	17	65.4%					
Parent Highest Education										$\chi^2(2) = 1.066, p > .05$
	<i>Intermediate vocational ('MBO')</i>	5	19.3%	3	11.5%					
	<i>Higher vocational ('HBO')</i>	7	26.9%	10	38.5%					
	<i>University ('WO')</i>	14	53.8%	13	50%					
N-CDI 3		26		86.91 (21.48)	26		94.12 (6.67)			$t(29.775) = -1.633, p > .05$

2.2 Materials

2.2.1 Speech materials

Both for Cantonese and Mandarin, a stimulus set of 16 minimal pairs was used for each phonological contrast (affricate contrast and tone contrast). The materials used for the perceptual training, the perception tests, and the production tests were subsets of those stimuli. For the Cantonese materials, see section 2.2.1.1 and Appendix B (Table B1), and for the Mandarin materials, see section 2.2.1.2 and Appendix B (Table B2).

All stimuli were disyllabic pseudowords that were phonotactically legal in Cantonese and Mandarin (as well as Dutch). The monosyllables that they consist of – like all phonotactically legal monosyllables in Cantonese and Mandarin Chinese – are existing words and are connected to Chinese characters (for Cantonese: Kwan, Tang, Chiu, Wong, Wong, & Zhong, 2006; for Mandarin: *Xinhua Dictionary*, 2006).

In addition, Dutch speech materials were used for instructions, encouragement, and practice (see section 2.2.1.3).

2.2.1.1 Cantonese

For the alveolar affricate contrast stimuli, the first syllable was always the open central unrounded vowel [a]. The second syllable started with the crucial affricate (i.e., either the unaspirated [ts] or the aspirated [ts^h]). The rest of the second syllable consisted of either a vowel (V), a vowel plus nasal (VN), a glide plus vowel (GlideV), or a glide plus vowel plus nasal (GlideVN). Both syllables carried tone 1 (High-Level).

For the tone contrast stimuli, the crucial tone (i.e., either tone 2 (High-Rising) or tone 5 (Low-Rising)) occurred on the second syllable. As in the affricate contrast stimuli, the first syllable was again the open central unrounded vowel [a] with tone 1 (High-Level), and the second syllable had a structure of CV, CVN, CGlideV, or CGlideVN again. The onset consonant could either be a plosive, fricative, nasal, or central or lateral approximants. No affricates were used in the onset, as this might affect the perception of the target contrast in the affricate contrast stimuli.

Three female native Cantonese speakers (*aged 21, 26, and 28 years*) from Guangzhou city in China were recorded. They read all 64 stimuli (i.e., 32 stimuli (16 minimal pairs) per contrast) in a random order, and in a clear citation style. All recordings were made onto a computer using Adobe audition software (Sample Rate: 44.1 kHz, Channels: Stereo, and Resolution: 16-bit) in a sound-proof booth and were afterwards segmented with the software package PRAAT (Boersma & Weenink, 2001). All stimuli were recorded multiple times; 11 tokens per speaker of each stimulus were selected for use in the training, and in the perception and production tests, as described in sections 2.4.1.1 (training), 3.1.1.2 (perception) and 4.1.1.2 (production).

2.2.1.2 Mandarin

The stimuli for the Mandarin contrasts were structured as described for the Cantonese contrasts (see 2.2.1.1; but note that the stimuli themselves were not identical). The crucial contrasts were now the unaspirated [tʂ] versus aspirated [tʂ^h] retroflex affricates, and Mandarin tone 2 (High-Rising) versus Mandarin tone 3 (Low-Dipping)).

Three female native Mandarin speakers (*aged 21, 26, and 29 years*) from the Beijing area were recorded for the stimuli. Following the same procedures in section 2.2.1.1, 64 stimuli (for each stimulus, 11 tokens per speaker) were selected. For details, see sections 2.4.1.1 (training), 3.2.1.2 (perception test), and 4.2.1.2 (production test)).

2.2.1.3 Dutch (instructional materials)

Several fragments were recorded by a female Dutch speaker (*age = 35 years*) in a child-directed speech style to serve as instructions at several points in the sessions (see section 2.3.2 for details).

Further, as practice stimuli, a set of 10 minimal pairs of disyllabic Dutch pseudowords was recorded (see Appendix B3). The first syllable was the open back unrounded vowel [a]. The second syllable contrasted the initial consonants [f] and [t], for instance, [a'fa:n] versus [a'ta:n]. These obstruents exist both in Dutch and Chinese, and were expected to be easy to distinguish for the participants.

In addition to the speaker who read the instructions, another two female native Dutch speakers (*aged 27 and 34 years*) were recorded for the practice stimuli, in the same way as

described for the Cantonese and Mandarin materials. They all read the stimuli in a clear citation style. Stimuli were recorded in a random order for multiple times. For each speaker, one token of each stimulus was selected, with a total of 20 stimuli per speaker.

2.2.2 Video materials

Because of the young age of the participants and the length of the experimental sessions (see Procedure section below), experiments were delivered in the form of animated video games. The video materials were newly constructed for the current study. Videos of two animal families were created with Adobe Photoshop and Flash CS 5, namely a dinosaur family and a panda family, each consisting of a mother figure and two baby figures. The dinosaur family was used for instructions, practice, and the affricate contrast, and the panda family for the tone contrast; both species were also used in the vocabulary study (dinosaurs for instructions and practice, and pandas for tests). In order to keep the videos engaging, different background pictures, accessories, and colors of the animals' clothes were used. For details of the videos and games, which were different for all the tests and training, see the Method sections of the corresponding chapters.

2.3 Procedure

2.3.1 Administration of the project

The current study was carried out during four sessions over a two-week period. There was a two or three days' gap between every two sessions (in order to allow some time to the participants to consolidate what they had learned from the previous session). The testing schedule of the four sessions was thus as shown in Table 3.

Participants were tested before, during, and after phonological perceptual training. The phonological perceptual training consisted of ten blocks, training the perception of the affricate and tone contrasts of the birth language. Each perceptual training block consisted of two parts: one part for the affricate contrast and one part for the tone contrast. The affricate contrast was presented first, followed by the tone contrast. The test was conducted at three instances, namely pretest, intermediate test and posttest, each testing the perception and the production of the phonological contrasts, and Chinese vocabulary. Each test instance contained two blocks: one block for the phonological contrasts and one block for Chinese

vocabulary. Each block for the phonological contrasts had four parts: one part for the affricate perception, one part for the affricate production, one part for the tone perception, and one part for the tone production. Affricate perception was tested first, affricate production second, tone perception third, and tone production fourth/last. Each block of the vocabulary test was conducted at the end of the experimental sessions where the perception and production tests were delivered, i.e., the first, the second, and the fourth (i.e., last) session.

As Table 3 describes, the first experimental session consisted of a block of general instructions and practice, a pretest block of perception and production of the phonological contrasts, two perceptual training blocks, and a pretest block of Chinese vocabulary; the second session contained two perceptual training blocks, an intermediate test block of perception and production of the phonological contrasts, and an intermediate test block of Chinese vocabulary; the third session was composed of four perceptual training blocks; and the fourth session had two perceptual training blocks, a posttest block of perception and production of the phonological contrasts, and a posttest block of Chinese vocabulary. There was a short break usually around 10-15 minutes after each block; thus each session lasted around an hour and a half, with a total of 6 hours for the complete experiment.

Table 3
Contents of the four experimental sessions

Session	Experimental content		Time (minutes)
I (Total duration: 90-110 minutes)	Block 1-General instructions & Practice		8
	Block 2-Pretest	Affricate Perception	4-5
		Affricate Production	2
		Tone Perception	4-5
		Tone Production	2
	Block 3-Perceptual training 1	Part 1-Affricates	6-8
		Part 2-Tones	6-8
	Block 4-Perceptual training 2	Part 1-Affricates	6-8
		Part 2-Tones	6-8
Block 5-Vocabulary pretest		5	
II (Total duration: 70-90 minutes)	Block 1- Perceptual training 3	Short practice	3
		Part 1-Affricates	6-8
		Part 2-Tones	6-8
	Block 2- Perceptual training 4	Part 1-Affricates	6-8
		Part 2-Tones	6-8
	Block 3- Intermediate test	Affricate Perception	4-5
		Affricate Production	2
		Tone Perception	4-5
		Tone Production	2
	Block 4- Vocabulary intermediate test		3
III (Total duration: 80-100 minutes)	Block 1- Perceptual training 5	Short practice	3
		Part 1-Affricates	6-8
		Part 2-Tones	6-8
	Block 2- Perceptual training 6	Part 1-Affricates	6-8
		Part 2-Tones	6-8
	Block 3- Perceptual training 7	Part 1-Affricates	6-8
		Part 2-Tones	6-8
	Block 4- Perceptual training 8	Part 1-Affricates	6-8
		Part 2-Tones	6-8

IV (Total duration: 70-90 minutes)	Block 1- Perceptual training 9	Short practice	3
		Part 1-Affricates	6-8
		Part 2-Tones	6-8
	Block 2- Perceptual training 10	Part 1-Affricates	6-8
		Part 2-Tones	6-8
	Block 3- Perception posttest	Affricate Perception	4-5
		Affricate Production	2
		Tone Perception	4-5
		Tone Production	2
	Block 4- Vocabulary posttest		

Note. Session 1 + 2 days' gap + Session 2 + 2 days' gap + Session 3 + 3 days' gap + Session 4

Participants were tested in a quiet room at their homes, to make the experience as comfortable as possible for them. They were seated together with the experimenter(s) during the experiments. Experimenter(s) gave instructions and encouragements to the participants as and whenever necessary¹⁵.

The experiment was delivered using Presentation software (version 14.7) from Neurobehavioral Systems by an experimental laptop (HP EliteBook 8540P with resolution 1366 x 768 pixels). Participants were seated at a comfortable viewing distance from the laptop screen. They heard the auditory stimuli through high-quality headphones (Sennheiser HD 280 64 ohm). A horizontally oriented button box with two buttons (MPI Dual Button box: Serial port via USB, Baudrate-38400, Data-8bit, and StopBit-1) was used. For the production data, participants were recorded through an audio recorder (Roland EDIROL R-09).

2.3.2 Instructions, task familiarization, and encouragement

The first session started with an introduction block in Dutch, in the shape of a video game similar the ones used in the test, explaining to the participants that they were going to play a game, and familiarizing them with the response buttons and with the task for the first perception test. During this block participants performed a task similar to that in the perception test but now involving an easy to distinguish Dutch contrast (see Appendix B3) to make sure that participants had fully acquired the task demands, and containing additional

¹⁵ For instance, participants touched or pointed at the computer screen to indicate their response, they were reinstructed to give their response by pressing a button on the button. Every time when participants finished a block, experimenter(s) cheered and encouraged the participants.

instructions to help them perform the task (i.e., if an incorrect response was given, the instructions told participants that they made a mistake and which button they should press; only when a correct response was received, the next trial started).

Similarly, the other sessions started with a similar but much shorter Dutch introduction. Further, there were short introduction parts during the sessions when a change of task occurred. Such Dutch introductions concerned perception, production, and vocabulary tests, as well as training.

Finally, throughout the sessions, in order to motivate participants between parts and blocks, there were encouragement videos containing recordings from the same Dutch speaker who read the instructions, also in a child-directed speech style (for example: "Well done, let's do it one more time." in Dutch).

2.4 Training: Method and Results

Perceptual training was used to explore the relearning advantage in the Chinese adoptees. Each training block consisted of 48 trials (24 per contrast). The same method was used for both the Cantonese and the Mandarin training.

2.4.1 Materials

2.4.1.1 Speech materials

For each phonological contrast, 12 out of 16 minimal pairs¹⁶ were used (24 stimuli per contrast). For each of the 24 stimuli per contrast, three tokens were used (one per speaker), giving 72 tokens per contrast per training block. In total, 720 tokens per contrast (24 stimuli per contrast * three tokens * 10 training blocks) were presented.

2.4.1.2 Video materials

There were 40 task videos, 40 positive and 40 negative feedback videos. 20 of the task videos showed three dinosaurs, and the other 20 showed three pandas. Similarly, half of the feedback videos showed three dinosaurs, and the other half three pandas. Each task video was combined with one positive feedback video and one negative feedback video, featuring the

¹⁶ In total, 16 minimal pairs for each phonological contrast were used in the present project, and 12 of them were presented in the training session.

same animals, with unique accessories and against a unique background. Each video showed the three animals standing next to each other, the mother animal in the middle with one baby animal on each side.

Each task video (see Figure 5) showed the animals speaking, each in turn. In each speaking turn, the animal was shown to open and close its mouth twice, for a total duration of 1,400 ms. Each task video started with 1,000 ms during which the animals were not moving; next the mother animal was shown to speak, followed by 1,500 ms without movement, then the baby on the left side of the screen was shown to speak, followed by 1000 ms without movement, and finally the baby on the right side of the screen was shown to speak. In total, each task video lasted 7,700 ms.

In each feedback video, one of the baby animals was moving and showing animation, while the other two animals were not. In a positive feedback video, either the right or the left baby animal was shown to cheer and laugh with colorful stars around the head (see Figure 6). In a negative feedback video, either the right or the left baby animal was shown to cry and rub their eyes or face (see Figure 7). Each feedback video lasted 1,000 ms.

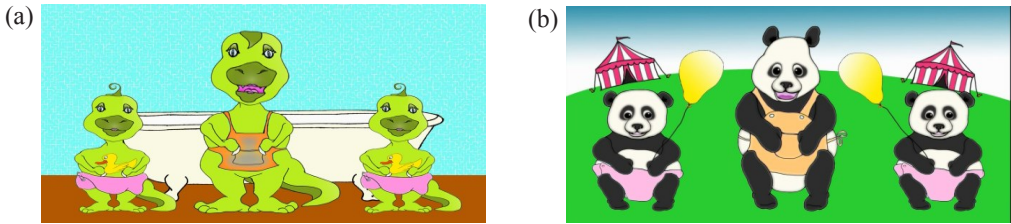


Figure 5. Example of a task video with the dinosaur family (a), and with the panda family (b).

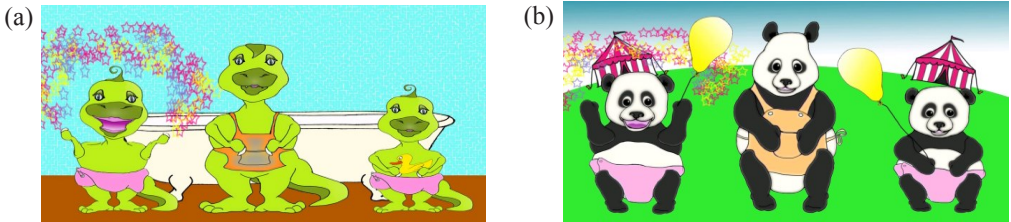


Figure 6. Example of a positive feedback video with the dinosaur family (a), and with the panda family (b).

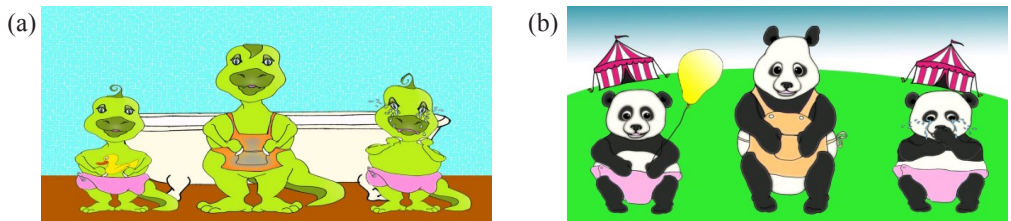


Figure 7. Example of a negative feedback video with the dinosaur family (a), and with the panda family (b).

2.4.2 Procedure

Participants were trained with an XAB discrimination task. In each trial, they heard three stimuli, two of which were the same, and their task was to judge whether the second or the third stimulus (i.e., A or B) was identical to the first stimulus (X). Each of the two affricates and each of the two tones occurred as X in the XAB task equally often. There were 24 trials per contrast in each training block, all with unique items¹⁷.

Each trial contained one stimulus from each of the three speakers, in a fixed order. Each speaker was paired with one of the three animals in the video. In terms of the video game, each trial started with the mother animal pronouncing one word (X), followed by first the baby animal on the left side of the screen (A), and second the baby animal on the right side of the screen (B).

¹⁷ Note that the deviating stimulus in one trial corresponded with X in another trial.

In each trial, the three audio stimuli were presented in the order XAB, aligned with the task video such that one animal was shown to speak during the presentation of each auditory stimulus. The first auditory stimulus (X) was presented at 1,000 ms after onset of the trial. There was an inter-stimulus interval (ISI) of 1,500 ms between the offset of the first and the onset of the second audio stimulus (i.e., X and A), and an ISI of 1,000 ms between the offset of the second and the onset of the third audio stimulus (i.e., A and B). Previous work suggests that ISIs of this length motivate phonological processing, in contrast to shorter ISIs which stimulate phonetic processing (Werker & Logan, 1985). Up to this point the trial lasted 7,700 ms. Participants responded by pressing one of the two response buttons, the left button for the left baby animal (A), or the right button for the right baby animal (B). Responses were registered from the offset of stimulus B. There was no time limit for responses, and the last frame of the task video stayed on the screen until a response button was pressed.

Immediately after each correct response a positive feedback video was played, and immediately after each incorrect response a negative feedback video was played. The feedback video was chosen such that the baby animal associated with the intended response displayed the feedback movement. As there was no break between the task video and the feedback video, and the background and accessories were the same, there was no transition between the task and feedback videos noticeable to the participants.

2.4.3 Results

The results of the training are presented in Figures 8 and 9. There were no unexpected results in the perceptual training for either group. Participants' performance fluctuated across all training blocks, and did not change significantly. There was no difference between the Chinese adoptees and the Dutch controls. The Chinese controls outperformed the adoptees and the Dutch controls. The high-accuracy performance of the Chinese controls suggests that the task was suitable for the children in the age range that was investigated in the present project, and it also confirms the quality of the Chinese materials. For details of analyses, please see Appendix C.

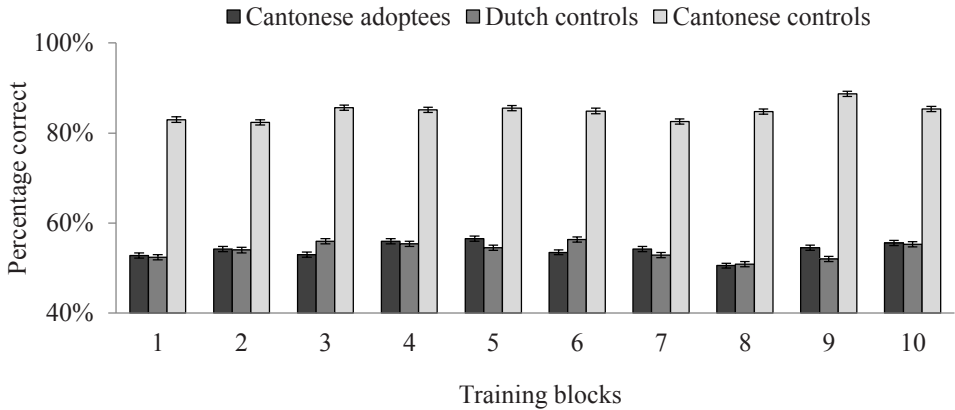


Figure 8. Percentage correct for the Cantonese adoptees, the non-adopted Dutch controls, and the non-adopted Cantonese controls (affricates and tones collapsed). Error bars represent standard errors.

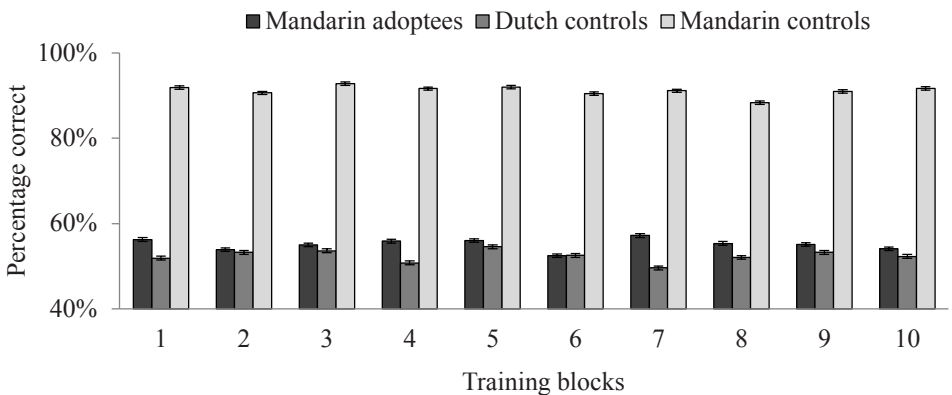


Figure 9. Percentage correct for the Mandarin adoptees, the non-adopted Dutch controls, and the non-adopted Cantonese controls (affricates and tones collapsed). Error bars represent standard errors.

2.5 Questionnaire on participants' characteristics

For all participants, a questionnaire was used (see Appendix D) to collect information on age, gender, language background, education, music skills, and parents' educational levels. In addition, for the adoptees and the Dutch controls, the questionnaire also asked whether or not they have siblings who were adopted from China, and whether they had visited China, etc. Particularly for the Chinese adoptees, this questionnaire (see Appendix D2) also included several specific questions about their adoption history, such as the age of adoption, length of residence in the Netherlands, birth language background, birth language fluency at the time

when they were adopted and when they were tested, and birth language re-exposure after adoption, etc. The questionnaire was completed by the participants' parents.

From the questionnaire, six control variables were summarized and assessed in the Participants Section 2.1. They were Age, Gender, Siblings, China Visits, Music, and Parent Highest Education.

Chapter 3: Perception of the phonological contrasts of the birth language

This chapter explores the perception of birth language phonological contrasts by the adopted Chinese children, as compared to non-adopted Dutch and Chinese children. The main research question addressed is: 1) do the adopted Chinese children perceive their birth language phonological contrasts more accurately than the Dutch control children, before and after the training? A secondary research question is: 2) do the Chinese control children perceive their native language contrasts at a high level of accuracy, and more accurately than the adopted Chinese children and the Dutch controls? This second question assesses the feasibility of the task for children of this age group (i.e., can children in this age range perform the task well when it is in their native language), as well as the quality of the stimuli (i.e. do they consist of accurate Cantonese recordings). Two experiments were done to investigate these questions. Experiment 1 tested the Cantonese adoptees, one group of Dutch controls, and the Cantonese controls, and Experiment 2 the Mandarin adoptees, another group of Dutch controls, and the Mandarin controls.

3.1 Experiment 1: Cantonese perception

This experiment tested the perception of the Cantonese affricate and tone contrasts. As described in Chapter 2, each test contained a part testing the affricates and a part testing the tones.

3.1.1 Method

3.1.1.1 Participants

Participants were the 22 Cantonese adoptees, 23 Dutch controls, and 22 Cantonese controls described in Chapter 2 (section 2.1.1).

3.1.1.2 Materials

3.1.1.2.1 Speech materials

In each of the three perception tests, both for the affricate and the tone contrasts, all 16 minimal pairs were used. For each of the 32 stimuli per contrast, three tokens were used (one from each speaker), with a total of 96 tokens per contrast.

3.1.1.2.2 Video materials

There were 12 test videos, and 12 series of eight motivation videos. The test videos were similar to the task videos used in the training (see section 2.4.1.2), showing three animals from the same family speaking, each in turn. Each video showed a unique background and unique clothing and accessories for the animals. The size of the test videos was horizontally compressed to 2/3 of the screen (Figure 10).

The motivation videos showed a single baby dinosaur or baby panda jumping up a eight-step stairway towards a gift box at the top of the screen. The motivation videos formed a series of eight videos, with the baby animal starting at the bottom of the stairs and jumping one step further in each consecutive video. The final motivation video of each series showed the baby animal reaching the top of the stairs, the gift box opening, and a unique gift popping out, accompanied by stars, a balloon and cheerful sound effects (Figure 11). Each motivation video lasted 1,000 ms. The motivation videos took up 1/3 of the screen horizontally, such that they could be presented next to the test videos, either on the left or on the right side of the screen (Figure 10).

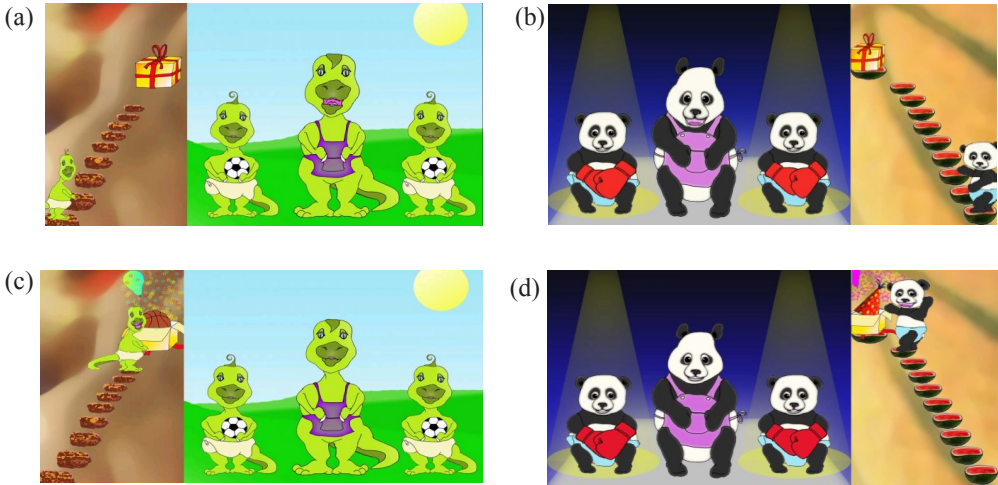


Figure 10 (a, b, c, d). Examples of video stills from a test trial, each containing a test video and a motivation video.

3.1.1.3 Design and procedure

The task used was similar to that in the training, except that now there was no feedback about the correctness of the responses. As in the training, an XAB discrimination task was used: each trial contained three auditory stimuli, two of which were the same, and participants were asked to judge whether the second or the third stimulus (i.e., A or B) was identical to the first stimulus (X). The task was again presented as a video game. Sound and video materials were aligned in the same way as in the training (see section 2.4.2 for details).

In each test, affricates were tested first, followed by the tones, in separate parts. Each part contained 16 trials and 16 minimal pairs. Each of the target affricates and tones occurred as X in the XAB task equally often. The 16 trials were grouped in two series of eight trials. Each series of eight trials used a single test video of the animals (i.e., keeping background and animals' accessories the same), and a series of eight motivation videos showing a single baby animal (from the same species as in the test video) jumping up the stairs from bottom to top. Between the series of eight trials, the position of the motivation videos changed from the left side to the right side of the screen. The test and motivation videos were simultaneously shown on the screen (Figure 10), but were played in consecutive order.

Each trial started with the test part, which was similar to the training trials up to the point that a response was given. Note that in contrast to the training, there was no feedback about the correctness of the response; therefore, the motivation video was played, to indicate to the participant that a response was received and to stimulate them to carry on with the experiment. In terms of the video game, each series of eight trials formed a coherent story line, by means of the eight motivation videos showing the baby animal climbing the stairs and being rewarded with a gift.

3.1.2 Results and discussion

Based on the reaction time (RT) distribution, 51 responses (1.2%) with an RT longer than 10,000 ms¹⁸ were considered as outliers and excluded from analysis. Mauchly's Test of Sphericity was used to assess whether the assumption of Sphericity was violated, and if so, degrees of freedom were corrected using Greenhouse-Geisser estimates of Sphericity.

3.1.2.1 Adoptees vs. Dutch controls

Figure 11 shows a crucial difference between the Cantonese adoptees and the Dutch controls: both groups perceived the contrasts similarly in the pretest and the intermediate test, but in the posttest the adoptees outperformed the controls. Whereas the Dutch controls did not improve over time, the Cantonese adoptees showed improved performance in the posttest.

¹⁸ Note that while RTs of 10,000 ms are rather long for adult studies, a cut-off of 10,000 ms was deemed appropriate for the present study, because children tend to respond relatively slowly. E.g., the children in the current study habitually touched or pointed to the computer screen to indicate their responses; only after the experimenter(s) instructed them, they would respond appropriately by pressing the response button.

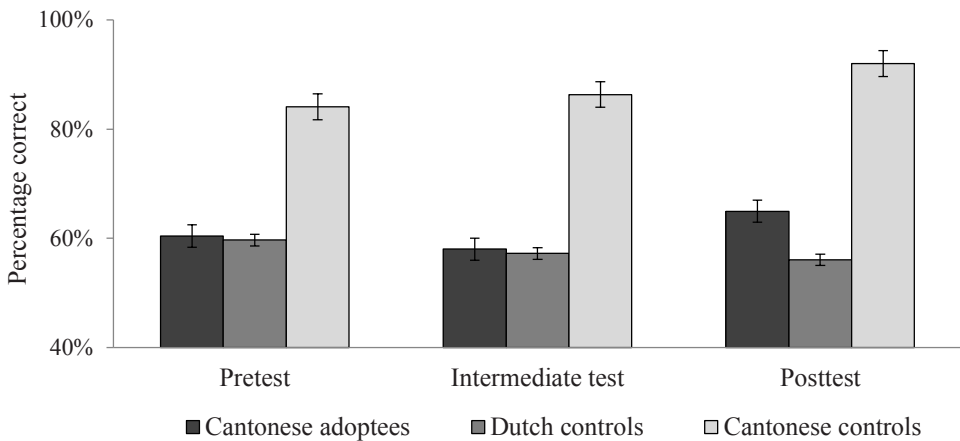


Figure 11. Percentage correct for the Cantonese adoptees, the non-adopted Dutch controls, and the non-adopted Cantonese controls (affricates and tones collapsed). Error bars represent standard errors.

To compare the adoptees against the Dutch controls, Analyses of Variance (ANOVAs) —by subjects (F_1) and by items (F_2)— were carried out with the proportions of correct responses as dependent variable, using the following independent variables¹⁹: Group (Cantonese adoptees and Dutch controls), Test instance (Pretest, Intermediate test, and Posttest), Contrast type (Affricates and Tones). The Parent Highest Education was used as a covariate, because this was the only control variable that was significantly different between the Cantonese adoptees and Dutch controls (see section 2.1.1).

Confirming what Figure 11 shows, there was a significant interaction between Group and Test instance, $F_1(2, 84) = 5.462, p < .01, \eta^2_p = .115, F_2(2, 60) = 3.253, p < .05, \eta^2_p = .098$. Following up on the interaction, the effect of Group was assessed at each test instance; there was no group difference in the pretest, $F_1(1, 43) = .178, p > .05, \eta^2_p = .004, F_2(1, 30) = .083, p > .05, \eta^2_p = .003$, or in the intermediate test, $F_1(1, 43) = .111, p > .05, \eta^2_p = .003, F_2(1, 30) = .138, p > .05, \eta^2_p = .005$, but indeed, crucially there was a significant group difference in the posttest, $F_1(1, 43) = 8.160, p < .01, \eta^2_p = .159, F_2(1, 30) = 9.701, p < .01, \eta^2_p = .244$, with a higher proportion of correct responses for the Cantonese adoptees than for the Dutch controls.

¹⁹ Group was a between-subjects and within-items variable, Test blocks a within-subject and within-items variable, Contrast type a within- subjects and between- items variable, and the Highest Educational Level of Parent(s) a covariate.

Further, still following up on the interaction between Group and Test instance, the effect of Test instance was assessed for each Group. As Figure 11 illustrates, there was no difference between Test instances²⁰ for the Dutch controls, $F_1(1.509, 33.208) = 2.251$, $p > .05$, $\eta_p^2 = .093$, $F_2(1.429, 42.881) = .581$, $p > .05$, $\eta_p^2 = .019$, but for the Cantonese adoptees, on the other hand, there was a significant effect of Test instance, $F_1(2, 42) = 4.820$, $p < .05$, $\eta_p^2 = .187$, $F_2(2, 60) = 3.225$, $p < .05$, $\eta_p^2 = .097$. Subsequent analyses show that the Cantonese adoptees performed significantly better in the posttest in comparison with the intermediate test, $F_1(1, 21) = 7.845$, $p < .05$, $\eta_p^2 = .272$, $F_2(1, 30) = 7.872$, $p < .01$, $\eta_p^2 = .208$. There was no difference between the intermediate test and the pretest, $F_1(1, 21) = 1.258$, $p > .05$, $\eta_p^2 = .057$, $F_2(1, 30) = .576$, $p > .05$, $\eta_p^2 = .019$ ²¹.

There was no main effect of or interaction with Contrast type or Parent Highest Education²². This indicates that the two types of phonological contrasts (i.e., Affricates and Tones) were perceived similarly by the adoptees and the Dutch controls. Therefore, Affricates and Tones were collapsed in Figure 11; for a separate figure of each contrast, see Appendix E. Importantly, a similar F_1 ANOVA to the one used in the main analysis but including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates largely confirmed the main findings described above²³.

To sum up, the Cantonese adoptees demonstrated an advantage over the Dutch controls in perceiving their birth language contrasts, but only after finishing the perceptual training. In the pretest and the intermediate test, on the other hand, the adoptees showed a lack of sensitivity to the Cantonese affricate and tone contrasts, as their performance did not differ from that of the Dutch controls. Whereas the performance of the adoptees improved significantly over time, that of the Dutch controls did not. The present findings suggest that the phonological exposure received by the adoptees in their early childhood, even though limited to the first years of life, still led to a benefit when relearning the sounds of the language several years after adoption.

²⁰ Mauchly's Test indicated that the assumption of Sphericity for the main effect of Test moment was violated in both the F_1 ($\chi^2(2) = 8.254$, $p = .016$) and F_2 analysis ($\chi^2(2) = 14.777$, $p = .001$). Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of Sphericity, F_1 , $\epsilon = .755$, F_2 , $\epsilon = .715$.

²¹ Note that there was no significant difference between the posttest and the pretest, $F_1(1, 21) = 5.371$, $p < .05$, $\eta_p^2 = .204$, $F_2(1, 30) = 2.990$, $p = .094$, $\eta_p^2 = .091$.

²² Because there was no effect of the Highest Educational Level of the Parent(s), it was no longer included as a covariate in the follow-up analyses.

²³ Now, however, the difference between the intermediate test and the posttest was no longer significant. This might be due to the lack of statistical power when covariates were included. Additionally, there were significant effects of the covariates Age ($F_1(1, 37) = 5.611$, $p < .05$, $\eta_p^2 = .132$) and Music ($F_1(1, 37) = 4.196$, $p < .05$, $\eta_p^2 = .102$).

3.1.2.2 Adoptees and Dutch controls vs. Cantonese controls

Figure 11 also shows the results of the Cantonese controls; the Cantonese controls performed very well at all three test instances. ANOVAs similar to those described above but including the Cantonese controls showed that the Cantonese controls performed significantly better than the Cantonese adoptees as well as the Dutch controls (Cantonese controls versus Cantonese adoptees: $F_1(1, 42) = 78.141, p < .001, \eta^2_p = .650, F_2(1, 30) = 147.948, p < .001, \eta^2_p = .831$; Cantonese controls versus Dutch controls: $F_1(1, 43) = 126.654, p < .001, \eta^2_p = .747, F_2(1, 30) = 142.756, p < .001, \eta^2_p = .826$).

The high-accuracy performance of the Cantonese controls confirms that the task was feasible for children in this age range (which matched the range of the Cantonese adoptees and Dutch controls), as well as confirming the quality of the Cantonese materials.

3.1.2.3 Relationships among perception and Age, Age of Adoption, and Length of Residence

The Cantonese adoptees fell in a wide range of age at the time of testing (i.e., between 4;4 and 10;10 years old), as well as of age at the time of adoption (between 0;9 and 4;6 years) and time they had spent in the Netherlands (between 1;5 and 9;11 years). It was investigated whether Age, Age of Adoption (AoA), and Length of Residence (LoR) influenced perception of the birth language contrasts of the Cantonese adoptees. In order to do that, Kendall's Tau²⁴ correlation analysis was conducted between the perception and Age, AoA, and LoR (see Appendix F.1 for the statistical analyses). For this purpose, six different measures of performance were chosen: Accuracy Overall (averaging over Affricates and Tones, and all test instances), at the Pretest, at the Intermediate test, at the Posttest, for Affricates, and for Tones (averaging over all test instances). For the three 'time variables', Age, AoA, and LoR are linearly related. Moreover, LoR is significantly correlated with Age and with AoA. Note that the linear relation among the three factors, and the correlation between LoR and Age, are inevitable (because $AoA + LoR = Age$).

Correlations were assessed between each of the six measures of performance and each of the time variables; no significant correlation was revealed. Next, partial correlations were

²⁴ Kendall's Tau correlation, which is not highly sensitive to outliers, was used because of the big individual differences in terms of Age, AoA, and LoR among the participants.

assessed between each of the six measures and Age and AoA, while controlling for LoR. Again, there was no significant correlation²⁵.

Finally, in order to disentangle the correlation between AoA and LoR, we divided the 22 adoptees into two subgroups such that AoA and LoR were not significantly correlated in either group²⁶ (see Table F.1.1d in Appendix F.1). Again, there was no significant correlation with AoA. (Note, however, that this might have been due to the small number of adoptees in each sub-group.)

Interestingly, in one of the sub-groups (with AoA > 1;7 years) there were significant positive correlations between several performance measures and Age: the adoptees who were older at the time of testing tended to show better performance (in particular after training). There were also significant positive correlations between some performance measures and LoR: adoptees who had been in the Netherlands longer tended to perform better. Because the correlation between Age and LoR cannot be disentangled, however, this counterintuitive effect of LoR might reflect the (intuitively more likely) effect of Age.

3.1.2.4 Effects of Siblings (having adopted Cantonese siblings) and China Visits (having visited the Cantonese areas in China)

As described in Chapter 2 (section 2.1.1), several children both in the groups of the adoptees and of the Dutch controls had one or more adopted Cantonese siblings (i.e., the factor ‘Siblings’). Similarly, in both groups several children had visited the Cantonese areas in China (for adoptees: between adoption and test) (i.e., ‘China Visits’). In order to investigate whether there is any effect of the two variables on the perception of the Cantonese contrasts, two analyses were done. First, an F_I ANOVA was carried out as in the main analyses described in section 3.1.2.1, now including Siblings and China Visits as covariates. There were no effects of or interactions with either of these factors.

Second, an F_I ANOVA was conducted like the one described in section 3.1.2.1, but leaving out all covariates, and adding Siblings and China Visits as independent variables. Results were the same as in the analysis in section 3.1.2.1., and in addition showed significant interactions between Group and Siblings, $F_I(1, 38) = 5.874, p < .05, \eta_p^2 = .134$, and between

²⁵ Due to the linear relation ($AoA + LoR = Age$), the outcomes of the correlations with each performance measure, after controlling for LoR, were identical for Age and for AoA.

²⁶ Note that LoR and Age are inevitably still significantly correlated in both sub-groups.

Group and China Visits, $F_1(1, 38) = 6.227, p < .05, \eta^2_p = .141$. When follow-up analyses on the interactions were conducted, significant effects of Siblings ($F_1(1, 20) = 8.474, p < .01, \eta^2_p = .298$) and China Visits ($F_1(1, 20) = 7.406, p < .05, \eta^2_p = .270$) were found only in the group of Dutch control children but, interestingly, not in the group of adoptees. In particular, the Dutch control children who had adopted Cantonese siblings performed better than those who did not. Contrary to expectation, the Dutch control children who had never visited the Cantonese areas in China performed *better* than those who had. The lack of an effect of Siblings and China Visits for the Cantonese adoptees, and the unexpected direction of the effect of China Visits for the Dutch controls, may result from the dominant position of Mandarin in China, as will be argued in the General Discussion (section 3.3).

3.1.2.5 Summary

To recapitulate, the results of Experiment 1 showed that the Cantonese adoptees were no different from the Dutch controls in the pretest and the intermediate test, but they did outperform the Dutch controls in the posttest. Whereas the Dutch controls did not improve over the course of perceptual training, the Cantonese adoptees significantly improved from the intermediate to the posttest. Overall, there were no effects of Age, AoA, and LoR. For one subgroup, effects of Age and LoR were found, showing that the adoptees who were older at the time of testing, and those who had lived in the Netherlands longer tended to perform better (particularly after training). Note that the number of Cantonese adoptees in the present experiment is rather small for a correlational analysis; future research with bigger sample sizes is needed to further explore the role of each of the time variables Age, AoA, and LoR. Finally, there were no effects of Siblings and China Visits for the adoptees (between adoption and test); for the Dutch controls, there was a positive effect of Siblings and an unexpected negative effect of China Visits, as will be further discussed below.

3.2 Experiment 2: Mandarin perception

This experiment tested the perception of the Mandarin affricate and tone contrasts. As in Experiment 1, each test consisted of a part testing the affricates and a part testing the tones.

3.2.1 Method

3.2.1.1 Participants

Twenty-six Mandarin Adoptees, 26 Dutch Controls, and 18 Mandarin Controls participated in the present experiment, as described in Chapter 2 (section 2.1.2).

3.2.1.2 Materials

3.2.1.2.1 Speech materials

Like Experiment 1, all 16 minimal pairs (32 stimuli) of both the affricate and the tone contrasts were tested in each perception test. In total, 96 stimuli per contrast (32 stimuli per contrast * three tokens (one per speaker)) were used.

3.2.1.2.2 Video materials

The same video materials as in Experiment 1 were used in the current experiment.

3.2.1.3 Design and procedure

The task, design and procedure were the same as in Experiment 1.

3.2.2 Results and discussion

Results with RTs longer than 13,000 ms (67 responses or 1.4 % out of total) were treated as outliers and discarded from analyses. Mauchly's Test of Sphericity showed that the assumption of Sphericity was not violated.

3.2.2.1 Adoptees vs. Dutch controls

Figure 12 also shows an essential difference between the adoptees (in this experiment the Mandarin adoptees) and the Dutch controls: the two groups of children did not differ from each other significantly in their perception of the Mandarin contrasts in the pretest and in the

intermediate test; however, in the posttest the adoptees outperformed the controls. Neither the Mandarin adoptees nor the Dutch controls improved significantly over time.

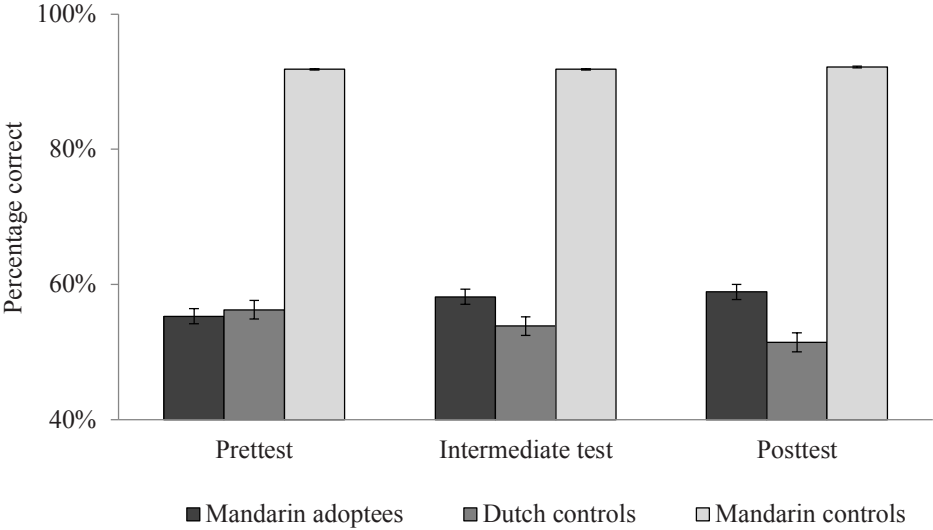


Figure 12. Percentage correct for the Mandarin adoptees, the non-adopted Dutch controls, and the non-adopted Mandarin controls (affricates and tones collapsed). Error bars represent standard errors.

For the comparison between the adoptees and the Dutch controls, the same ANOVAs as in Experiment 1 were performed. Again, the proportions of correct responses were used as dependent variable, Group (Mandarin adoptees and Dutch controls), Test instance (Prettest, Intermediate test, and Posttest), and Contrast type (Affricates and Tones) as independent variables. Note that in contrast to Experiment 1, there was no covariate included because the two groups of children were well-matched in all control variables (see section 2.1.2).

Results show a significant interaction between Group and Test instance, $F_1(2, 100) = 3.390, p < .05, \eta^2_p = .063, F_2(2, 60) = 4.240, p < .05, \eta^2_p = .124$. Follow-up analyses for the interaction were carried out to assess the effect of Group at each test instance and the effect of Test instance for each group. No group difference was found in the prettest, $F_1(1, 50) = .083, p > .05, \eta^2_p = .002, F_2(1, 30) = .074, p > .05, \eta^2_p = .002$; or in the intermediate test²⁷, $F_1(1, 50) = 2.909, p = .094, \eta^2_p = .055, F_2(1, 30) = 3.933, p = .057, \eta^2_p = .116$, but, importantly, there was a significant group difference in the posttest, with the Mandarin

²⁷ However, note that the group difference just missed significance in the intermediate test.

adoptees perceiving the Mandarin contrasts significantly better than the Dutch controls did: $F_1(1, 50) = 7.273, p < .05, \eta^2_p = .127, F_2(1, 30) = 9.767, p < .01, \eta^2_p = .246$.

There was no main effect of Test instance, either for the adoptees, $F_1(2, 50) = 1.350, p > .05, \eta^2_p = .051, F_2(2, 60) = .423, p > .05, \eta^2_p = .014$, or for the controls, $F_1(2, 50) = 2.224, p > .05, \eta^2_p = .082, F_2(2, 60) = .945, p > .05, \eta^2_p = .031$.

In line with Experiment 1, no effect of or interaction with Contrast type was found in Experiment 2, showing that the adoptees and the Dutch controls perceived the Mandarin affricate and tone contrasts in a similar way. The two types of contrasts thus again were collapsed in Figure 12; for a separate figure of each contrast, see Appendix E. Finally, an F_1 ANOVA similar to that in the main analysis but including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates largely showed the same results²⁸.

Confirming the finding of Experiment 1, Experiment 2 also showed an advantage for the (Mandarin) adoptees over the Dutch controls in perception of the Mandarin contrasts after completing the perceptual training (but no such group difference in the pretest or the intermediate test). In contrast however to Experiment 1, where the Cantonese adoptees did improve significantly from the intermediate test to the posttest, neither group of children improved their performance over time in Experiment 2. Another subtle difference between Experiments 1 and 2 is that the group difference between the Mandarin adoptees and the controls, although not reaching significance, is noticeable even in the intermediate test. These results thus provide additional evidence that phonological knowledge acquired during early childhood helps the adoptees relearn the sounds of their birth language later in life, even though they have been completely cut off from their birth language for several years.

3.2.2.2 Adoptees and Dutch controls vs. Mandarin controls

As also shown in Figure 12, the Mandarin controls perceived their native contrasts with high accuracy at each test instance. As in Experiment 1, ANOVAs including the Mandarin controls revealed that the Mandarin controls outperformed both the Mandarin adoptees and

²⁸ There was a significant effect of China Visits ($F_1(1, 44) = 8.164, p < .01, \eta^2_p = .157$), and several significant interactions with several covariates: namely, between Contrast type and Age ($F_1(1, 44) = 4.130, p < .05, \eta^2_p = .086$), between Contrast type and Siblings ($F_1(1, 44) = 4.699, p < .05, \eta^2_p = .096$), between Contrast type and Music ($F_1(1, 44) = 7.442, p < .01, \eta^2_p = .145$), between Contrast type and Parent Highest Education ($F_1(1, 44) = 4.657, p < .05, \eta^2_p = .096$), between Test moment and Music ($F_1(2, 88) = 3.537, p < .05, \eta^2_p = .074$), and among Test moment, Contrast type and Music ($F_1(2, 88) = 3.701, p < .05, \eta^2_p = .078$). Additionally, a significant effect of Contrast type (with higher scores for affricates than tones) was found in the F_1 analysis, but not confirmed in the F_2 analysis.

the Dutch controls (Mandarin controls versus Mandarin adoptees: $F_1(1, 42) = 293.980$, $p < .001$, $\eta^2_p = .875$, $F_2(1, 30) = 170.672$, $p < .001$, $\eta^2_p = .851$; Mandarin controls versus Dutch controls: $F_1(1, 42) = 433.376$, $p < .001$, $\eta^2_p = .912$, $F_2(1, 30) = 393.893$, $p < .001$, $\eta^2_p = .929$).

As in Experiment 1, the Mandarin controls' high-accuracy performance also confirms the feasibility of the task for the children in the age range tested in the present study, as well as the quality of the Mandarin materials.

3.2.2.3 Relationships among perception and Age, AoA, and LoR

Like the Cantonese adoptees in Experiment 1, the Mandarin adoptees had a varying age range (i.e., between 4;1 and 10;10 years old), age of adoption (0;10 to 5;8 years) and length of residence (from 0;3 to 9;10 years). Kendall's Tau correlation analysis was applied to investigate the effects of Age, AoA, and LoR on the perception of the birth language contrasts of the Mandarin adoptees (see Appendix F.1 for the statistical analyses). As in Experiment 1, six performance measures were used, namely, Accuracy Overall, at the Pretest, at the Intermediate test, at the Posttest, for Affricates, and for Tones. As before, the three 'time variables' Age, AoA, and LoR have a linear relation, which cannot be disentangled, and LoR and Age are inevitably correlated.

Following the procedure of Experiment 1, correlations between each of the six measures and each of the time variables were assessed first. The results of this assessment revealed significant positive correlations between several measures of performance and Age: adoptees who were older at the time of testing tended to perform better, particularly after training. There were also significant positive correlations between some performance measures and LoR: Adoptees who had been longer in the Netherlands tended to perform better. Again, given the inevitable correlation between Age and LoR, the LoR effect was assumed to be a reflection of the Age effect. There were no significant correlations with AoA.

Partial correlations between each of the six measures of performance and Age and AoA, with LoR controlled, largely showed the same correlations between performance measures and Age as described above: adoptees with an older age at the time of testing showed better performance, particularly after training. Due to the linear relation between Age and AoA, the two variables showed identical correlations.

Finally, the 26 adoptees were divided into two subgroups in an attempt to disentangle the AoA/LoR relationship: An early-adopted sub-group (AoA < 1;6 years) and a late-adopted sub-group²⁹ (AoA >= 1;6 years). Within each sub-group AoA and LoR were not significantly correlated³⁰ (see Table F.1.2d in Appendix F.1). Again, significant correlations between some measures of performance and Age were found, but only in the late-adopted group. In that group, adoptees who were older when tested performed better, particular after training (LoR demonstrated similar effects: adoptees who had been in the Netherlands longer performed better). Most interestingly, within the same sub-group, a significant positive correlation between performance in the pretest and AoA was found: Adoptees who were older at the time of adoption showed better performance in the pretest. This result suggests that the Mandarin adoptees who had been exposed to their birth language longer before adoption were likely to have preserved better phonological knowledge of their birth language, and thus were able to perceive the contrasts better before re-exposure. As for the other sub-group, there was no significant correlation of any of the performance measures with any of the time variables; note however that this might be due to the small number of adoptees in this sub-group.

3.2.2.4 Effects of Siblings and China Visits

Like the participants in Experiment 1, several Mandarin adoptees and Dutch controls had one or more adopted Mandarin siblings ('Siblings'), and some children in both groups had visited the Mandarin areas in China (for adoptees: between adoption and test) ('China Visits') (see section 2.1.2). Two analyses were carried out to investigate the effects of Siblings and China Visits on perception of the Mandarin contrasts. First, an F_1 ANOVA similar to that conducted in section 3.2.2.1, now including Siblings and China Visits as covariates, showed a significant interaction between Contrast type and Siblings, $F_1(1, 48) = 5.171, p < .05, \eta_p^2 = .097$. Follow-up analysis on the interaction revealed null effect of either Contrast type or Siblings. There was also a main effect of China Visits, $F_1(1, 48) = 6.778, p < .05, \eta_p^2 = .124$, showing that the participants who had visited the Mandarin areas in China (for adoptees: after adoption) performed better than those who did not.

²⁹ One adoptee adopted at the age of 68 months was excluded from the re-grouping procedure as an outlier, being much older at the time of adoption than the other children.

³⁰ Note that LoR and Age are inevitably still significantly correlated in both sub-groups.

Second, a similar F_I ANOVA to that in section 3.2.2.1, again adding Siblings and China Visits as independent variables, largely³¹ confirmed the results presented above: a significant interaction between Contrast type and Siblings, $F_I(1, 45) = 4.688, p < .05, \eta^2_p = .094$. Follow-up analyses on this interaction showed no effect of Siblings either for the affricate contrast or for the tone contrast, but did show a significant effect of Contrast type for the children with adopted siblings. The participants with adopted Mandarin siblings perceived the Mandarin tone contrast better than the Mandarin retroflex affricate contrast, ($F_I(1, 7) = 7.332, p < .05, \eta^2_p = .512$); those without adopted Mandarin siblings showed no difference between affricates and tones. Finally, there was also a main effect of China Visits, ($F_I(1, 45) = 4.822, p < .05, \eta^2_p = .097$), reflecting better performance in the children who had visited the Mandarin areas in China (for adoptees: after adoption) compared to those who did not.

3.2.2.5 Summary

The principal results of Experiment 2 are consistent with Experiment 1. The Mandarin adoptees demonstrated significantly better performance than the Dutch controls in the posttest, even though they did not perform differently from the control group either in the pretest or in the intermediate test. Further, neither group improved significantly across the three test instances (contrary to Experiment 1 where the adoptees but not the controls showed significant improvement over time). Overall, significant effects of Age and LoR were found in all adoptees, showing that older adoptees, and those who had been in the Netherlands longer, performed better (particularly after training) (in contrast to Experiment 1 where such effects were only present in a sub-group of 14 adoptees). Crucially, for the majority of the Mandarin adoptees who were adopted after the age of 1;6 years, there was a significant effect of AoA on the performance in the pretest: adoptees with a later AoA performed better before re-exposure. Further, the participants who had one or more adopted Mandarin siblings were better at perceiving the Mandarin tone contrasts. Finally, having visited the Mandarin areas in China led to better perception of the Mandarin contrasts (differing once again from Experiment 1).

³¹ Results show that the interaction between Group and Test moment was gone, in contrast to the main analysis in section 3.2.2.1., however, it did show a significant three-way interaction among Group, Test moment and Siblings, $F_I(2, 90) = 4.094, p < .05, \eta^2_p = .083$ (also in contrast to the main analysis).

3.3 General discussion

3.3.1 Long-term benefit of early phonological experience

Both experiments in the present study show a clear and consistent advantage for the Cantonese and the Mandarin adoptees over the Dutch controls in perception of the Chinese sounds after completing the perceptual training program. This finding is in line with previous studies on international adoptees who experienced a complete cut-off from their birth language for at least a decade (e.g., Choi, 2014; Pierce et al., 2014; Singh et al., 2011). The result also matches that reported in earlier studies on heritage language learners who were exposed to a language briefly during their early childhood and received minimal continued exposure later in life (e.g., Bowers et al., 2009; Oh et al., 2003; Tees & Werker, 1984). The results of the present study thus confirm and expand the insights from previous studies, by providing evidence for the existence of a long-term benefit of the phonological knowledge acquired in early childhood in young internationally adopted children.

Interestingly, before and between the perceptual training blocks, the Chinese adoptees did not yet perceive their birth language contrasts better than the Dutch controls did. This result agrees with that of Singh et al. (2011) and Choi (2014), in which teenaged and adult adoptees did not differ from the control participants prior to the training (i.e., in the pretest). Interestingly, this also accords with the behavioral result in Pierce et al (2014) who showed that at the neural level a well-preserved memory of the birth language tones existed in young Chinese adoptees. Further, it resembles earlier studies in adult adoptees which reported no evidence of phonological memory of the birth language without any training (Pallier et al., 2003; Ventureyra et al., 2004). Thus, the results from previous studies as well as from the present study confirm the impression commonly reported by international adoptees and their adoptive parents (see Chapter 1) that international adoptees simply forget their birth language some time after adoption, if access to the birth language is disrupted.

The adoptees' pattern of results is especially striking given that the two groups of Dutch control children in Experiments 1 and 2 showed a *falling* pattern of performance over time, albeit non-significantly. This pattern matches that of the control participants in Singh et al. (2011). One likely explanation for such a performance decrease might be in terms of the perceptual difficulty of the Chinese contrasts, as predicted by the Perceptual Assimilation

Model (Best & Tyler, 2007). Note also that the training program was limited in time and stimuli (i.e., contained only isolated pseudowords without conversational context) to maintain experimental control. Previous research suggests that a large amount of training, preferably in a linguistically varied context, is generally required for second language learners to learn non-native contrasts. For instance, Japanese learners of English were only able to learn to discriminate the English [r] versus [l] after receiving intensive training in a linguistically rich context over a long period of time (MacKain & Best, 1981). Despite the difficulty of the Chinese contrasts and the short training program in the current study, the Chinese adoptees improved their performance (although non-significantly in the case of the Mandarin adoptees), such that these results suggest the adoptees' early exposure to the Chinese sounds aids them to relearn the contrasts of their birth language, even several years after adoption.

3.3.2 Effects of AoA and Age (and LoR)

Within the groups of adopted children, there were some indications that AoA and Age might affect perception of the birth language contrasts. With respect to AoA, being adopted at a later age leads to better perception of the birth language contrasts *before* any re-exposure, based on the data of the majority of Mandarin adoptees. As for the Cantonese adoptees, however, no effect of AoA was found. This inconsistency may be attributed to the small number of Cantonese adoptees (overall and in either sub-group) compared to the Mandarin adoptees. Further, it might be the case that the Cantonese adoptees have received less input of Cantonese Chinese before adoption (due to the bilingual situation in their birth regions) in contrast to the Mandarin monolingual adoptees (see section 1.6). It is likely that the variations in the amount of Cantonese input may have washed out any possible effect of AoA. These results therefore should be interpreted with caution, because the complete groups of Cantonese and Mandarin adoptees failed to show consistent effects of AoA. As suggested by Oh et al. (2010) and Singh et al. (2011), a large sample size and a wide range of AoA are necessary for the investigation of the effect of AoA. The present study has a much bigger sample size and a much wider range of AoA than previous studies; however, the sample size is relatively small compared to the wide range of AoA, and small for a correlation analysis.

With respect to Age, the older the adoptees were at the time of testing, the better their performance generally was, for the Mandarin adoptees and less clearly so for the Cantonese adoptees. In contrast to the result for the Chinese adoptees, there was no effect of Age for the Dutch controls. This is striking because the non-adopted Dutch children, like the adoptees,

ranged widely in age (4-10 years). A possible explanation for the lack of an effect of Age in the controls is that there were two factors at play. On the one hand, the youngest children easily fell within the Critical Period for language learning (Johnson & Newport, 1989; Long, 1990), and might be expected to perform better in training. On the other hand, task used in this study may have favored the (cognitively more mature) older children. The advantage of a younger age with respect to language learning and the advantage of an older age with respect to cognitive development might then have cancelled each other out.

3.3.3 Effects of Siblings and China Visits

The effects of the variables Siblings and China Visits were different in Experiments 1 and 2. With respect to Siblings, in Experiment 1 (Cantonese), only for the Dutch control children, having adopted siblings was related to better performance. In Experiment 2 (Mandarin), having adopted siblings also had effects on the performance, but for both the Mandarin adoptees and the Dutch controls, and in a different way: Those participants with adopted Mandarin siblings perceived tones better than affricates, while those without such a sibling perceived tones and affricates in a similar pattern. It is unclear why having an adopted sibling is related to a differential perception of tones versus affricates in particular; however, the effect of having adopted siblings, in general, may stem from at least two sources: 1) the children are likely to have overheard Chinese speech sounds from their adopted siblings, which has facilitated their perception of the Chinese contrasts, and 2) having adopted siblings may have affected the participants' motivation.

With respect to China Visits, in Experiment 1 (Cantonese), whereas there was no effect for the adoptees, an unexpected negative effect was found for the control children. It is not clear why visiting China affected performance negatively for the Dutch controls, but see the discussion in section 3.3.3. In Experiment 2 (Mandarin), in contrast, there was a positive effect of China Visits both for the adoptees and the Dutch controls. This suggests that having visited the Mandarin areas in China facilitated the perception of the Mandarin contrasts for all participants alike. Two probable sources may have contributed to the *positive* effect of China Visits: 1) visiting the Mandarin areas in China provided the participants with Mandarin exposure, which helped the participants to perceive the Mandarin contrasts, and 2) visiting China may have increased their motivation.

For the Cantonese adoptees (in contrast to the Mandarin adoptees and the two groups of Dutch controls), there was no effect of Siblings and China Visits. That is striking because the number of children with adopted siblings and the number of children who had visited China were both much larger in this group than in the other three groups (Siblings: 9 out of 22 adoptees; China Visits: 10 out of 22 adoptees) (see Table 1 in Chapter 2 and Table 4 below). The null effect of Siblings and China Visits is possibly attributable to the bilingual situation (and the dominant position of Mandarin Chinese) in the Cantonese areas (and all over China), as further explained in the next section. Another possible cause, particularly for the non-effect of Siblings, is the small number of adoptees who are likely to have overheard Cantonese speech sounds from their adopted siblings. It is arguable that the participants with *younger* adopted siblings have a bigger chance than those with *older* adopted siblings to overhear Chinese sounds from adopted siblings, because the older adopted siblings may have stopped using (or may even have lost their global communicative skills in) their birth language at the time when the younger siblings were adopted or born (c.f., Nicoladis & Grabois, 2002)³². Although in both groups five children had younger adopted siblings (Table 4), in the group of Cantonese adoptees only two of the younger siblings were reported to have spoken any Cantonese (according to reports from the adoptive parents)³³, whereas in the group of the Dutch controls, all younger siblings were reported to have spoken Cantonese.

³² The average time elapsed between the adoption of two Chinese children in one family in our sample was two years. For international adoption, the average waiting time between the start of an adoption procedure and actual adoption is around 18 months, http://www.adoptie.nl/p/1563/veelgestelde_vragen_over_de_adoptieprocedure/.

³³ One of them was not able to speak at the time of adoption due to a cleft palate; the other one did not speak at the time of adoption because he was only thirteen months of age and had a severely delayed development in many domains.

Table 4

Descriptive information for Siblings in Experiments 1 and 2

Participants	Without adopted siblings	With adopted siblings	
		With younger adopted siblings	With older adopted siblings
Cantonese adoptees	13	5	4
Dutch controls (1)	18	5	0
Mandarin adoptees	22	3	1
Dutch controls (2)	20	5	1

Note. Dutch controls (1) were tested in the Cantonese study, while Dutch controls (2) in the Mandarin study.

3.3.4 Impacts of monolingual versus bilingual input

As introduced in Chapter 1, Mandarin Chinese is dominantly used in China; as a consequence, the Cantonese areas are bilingual in Mandarin and Cantonese. This may account for the different effects of Siblings and China Visits in Experiments 1 versus 2. With respect to Siblings, Cantonese adoptees could have overheard both Cantonese and Mandarin sounds from their adopted siblings, which may have resulted in relatively limited exposure to Cantonese, thus failing to elicit an effect of Siblings. With respect to China Visits, it is highly possible that the participants who visited the Cantonese areas in China were mainly (or even exclusively) exposed to Mandarin Chinese, rather than to Cantonese Chinese. Any Cantonese sounds that the participants may have been exposed to during their visits may, in consequence, have been insufficient to elicit any positive effect. This does not explain the negative effect of China Visits for the Dutch controls in Experiment 1; however, note that only four control children visited the Cantonese areas in China.

The wide use of Mandarin Chinese in the Cantonese areas may also have contributed to the different patterns of results in Experiments 1 and 2 (compare Figures 11 and 12): the Cantonese adoptees improved significantly between the intermediate and the posttest, whereas the Mandarin adoptees did *not* improve significantly at all, but showed only a non-significant tendency towards improvement from the pretest to the intermediate test. This earlier (but non-significant) tendency towards improvement might be a result of the monolingual Mandarin input before adoption; the Cantonese adoptees' later improvement might be a result of the bilingual input (Cantonese-Mandarin) before adoption.

The present study is the first to have investigated relearning of the phonological contrasts of the birth language by a large number of adopted children. Consistent results from two experiments demonstrated a robust relearning benefit in the adoptees over the Dutch controls, although there had been no difference between them in their initial performance. Our findings suggest that once learned phonological knowledge can be maintained in memory for a long time and made accessible by re-exposure. Additionally, the data suggest that longer exposure to the birth language before adoption might help the adoptees maintain a better phonological memory of their birth language. Finally, having (and possibly hearing Chinese spoken by) adopted Chinese siblings facilitated the perception of Chinese contrasts, for the adoptees as well as the Dutch controls, as did visiting China.

Chapter 4: Production of the phonological contrasts of the birth language

This study set out to investigate the production of the phonological contrasts in the birth language by the adopted Chinese children, in comparison with non-adopted Dutch control children. The main research question is: do the adopted Chinese children produce their birth language contrasts more accurately than the Dutch control children, before and after perceptual training? Like the perception study in Chapter 3, two experiments were undertaken. Experiment 1 assessed the production of the Cantonese adoptees and a group of Dutch controls, while Experiment 2 evaluated the production of the Mandarin adoptees and a separate group of Dutch controls.

4.1 Experiment 1: Cantonese production

Experiment 1 tested the production of the Cantonese affricate and tone contrasts. As in the Cantonese perception experiment in Chapter 3, the production of the affricates and the tones was tested separately in two parts (see Table 3 in Chapter 2).

4.1.1 Method

4.1.1.1 Participants

Participants were the same Cantonese adoptees and Dutch controls as those in the Cantonese perception study, except for one drop-out amongst the Dutch controls³⁴. This resulted in 22 adoptees and 22 controls.

4.1.1.2 Materials

4.1.1.2.1 Speech materials

For each production test, both for the affricate and the tone contrasts, two high-frequency³⁵ minimal pairs were selected out of the 16 which were described in section 2.2.1 (see Table 5).

³⁴ One Dutch control child (*Age* = 5;3 years) dropped out of the production experiment because he was too shy to say the Chinese sounds.

³⁵ The two high-frequency minimal pairs were chosen based on the suggestions from two native Cantonese speakers (*Age* = 30 and 23 years old).

Two identical tokens (from the same speaker) were used for each stimulus, with a total number of eight tokens for each contrast. Only the two speakers who were assigned to the two baby animals in the video game (see section 2.2.2) were used.

Table 5

Test stimuli of each experimental contrast in Experiment 1 with Chinese characters.

Alveolar affricate contrast		Tone contrast	
Unaspirated	Aspirated	Tone 2 (High-Rising)	Tone 5 (Low-Rising)
[atɕɛ] "阿遮"	[atɕ ^h ɛ] "阿车"	[atou] "阿土"	[atou] "阿肚"
[atɕuŋ] "阿装"	[atɕ ^h uŋ] "阿仓"	[asœŋ] "阿想"	[asœŋ] "阿上"

Note. The first syllable [a] "阿" in both contrasts, and the second syllable in the affricate contrast carried tone 1, while the second syllable in the tone contrast carried either tone 2 or tone 5.

4.1.1.2.2 Video materials

In total, six test videos were used (two videos at each test instance: one for the affricate contrast and one for the tone contrast). The test videos were similar to the task videos used for training (see section 2.4.1.2), showing three animals from the same family with a unique background, and with the mother and one of the baby animals holding a microphone (Figure 13).

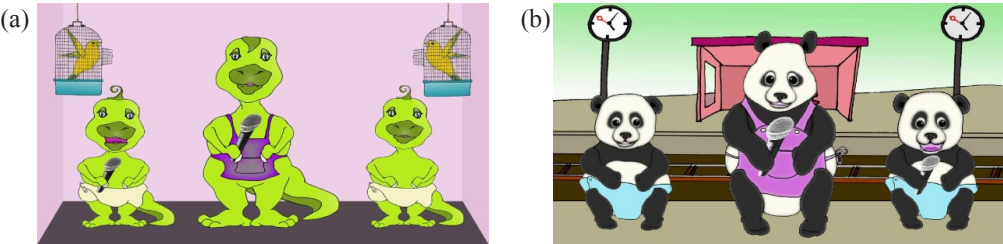


Figure 13 (a, b). Two examples of video stills from a test trial with dinosaurs and pandas.

4.1.1.3 Design and procedure

Participants were tested with an imitation task at three instances, i.e., before (pretest), in between (intermediate test), and after (posttest) perceptual training (see Table 3 in Chapter 2). Similar to the training and the perception experiments, the imitation task involved a video game, with sound and video materials aligned. Each trial started with the baby animal with

the microphone speaking a target word twice with 1,500 ms ISI³⁶. 1,000 ms later, the mother animal asked the participants (in Dutch) to repeat the target word twice, as accurately and clearly as possible.

As in the perception experiments, the affricate contrast was tested first, and then the tone contrast, in separate parts. In each part, four trials of two minimal pairs were used for practice (see Appendix B), and four trials of another two minimal pairs were used for test. The four practice trials were presented before test to familiarize the participants with the task. Participants were recorded individually in a quiet room in their homes with a high-quality audio recorder (Roland EDIROL R-09).

Eight target words for each contrast (with two tokens per word) were collected at each test for each child. Of the two tokens per target word collected from each participant, only the first token was used for the assessment experiment. The second token was used only if the sound quality of the first token was low. In total, 528 stimuli per contrast (4 target words * 3 test instances * 44 participants) were used in the assessment experiment.

4.1.1.4 Assessment of children's productions

4.1.1.4.1 Participants

Forty-five native Cantonese listeners participated in the assessment experiment. When tested, all participants were University students³⁷ from Guangzhou city in China. All had been living in the Guangzhou region since birth. They were randomly divided into two groups for two different types of assessment tasks (described in section 4.1.1.4.2) (see Table 6). They all received a small honorarium for their participation.

³⁶ Note that only the speakers who were assigned to two baby animals were used.

³⁷ Participants were recruited from five different Universities in Guangzhou city, namely, Jinan University, Sun Yat-sen University, South China Normal University, South China University of Technology, and Guangdong University of Foreign Studies.

Table 6

Descriptive information of the native Cantonese participants in the assessment experiment.

	Number	Mean age	Gender		Education	
Identification	24	22;8 years	12 F	12 M	12 BA	12 MA
Rating	21	22;3 years	11 F	10 M	10 BA	11 MA

Note. "F": Female, and "M": Male. The highest education on-going at the time of testing. "BA": Bachelor program, and "MA": Master program.

4.1.1.4.2 Materials

The productions that were recorded in Experiment 1 here formed the stimuli in the assessment experiment. For each of the two contrasts, 528 stimuli were used (see section 4.1.1.3).

Two different types of tasks were used to assess the children's pronunciation of the Cantonese contrasts, namely a two-alternative forced choice (2AFC) identification task and a rating task. In both tasks, participants were instructed to evaluate the pronunciation of the target affricates and tones only, respectively, and to ignore the rest of the phonological elements. Specifically, for the affricates they were asked to only assess the pronunciation of the onset consonant (i.e., the affricate) of the second syllable, while for the tones they were asked to only assess the pronunciation of the tone of the second syllable.

4.1.1.4.3 Procedure

The two experimental contrasts (i.e, affricates and tones) were tested in separate blocks on two consecutive days. One group of native Cantonese listeners participated in the identification test, and another group in the rating test.

In the identification test, participants were first presented with a minimal pair written in Chinese characters on the computer screen. The minimal pair was placed in two separate yellow rectangular boxes arranged horizontally in the middle of the computer screen, with a distinct space between the boxes. At 1,000 ms after the presentation of the visual display, participants heard a word played through their headphones. Their task was to identify the onset consonant of the second syllable (for the affricate contrast) or the tone of the second syllable (for the tone contrast) in the word they heard. To indicate their answer, they clicked on the matching word on the computer screen using a mouse.

In the rating test, participants were first presented with a single target word written in Chinese characters in the middle of the computer screen. 1,000 ms later, they heard a word played through their headphones. Their task was to evaluate the pronunciation of either the onset consonant of the second syllable (for the affricate contrast) or the tone of the second syllable (for the tone contrast). They had to rate how similar the pronunciation was to that of the word on the screen, on a four-point scale³⁸ (1 = completely different "完全不一样", 2 = a little similar "有点相似", 3 = very similar "非常相似", and 4 = identical "完全一样"). The four response options were displayed in four yellow rectangular boxes below the target word on the screen. To give a response, participants were asked to click on a box in the scale.

Each target word was played only once in the assessment experiment. Before each test, participants received eight practice trials to familiarize them with the task. They were allowed to adjust the audio volume to a comfortable level (with a predetermined minimal volume) during the experiment. The experiment was conducted in a quiet classroom at the participants' universities. Each test block took 25 - 28 minutes.

The experiment was delivered using the PRAAT program (Boersma & Weenink, 2001) on a laptop (HP EliteBook 8540P with resolution 1366 x 768 pixels). The auditory stimuli were played through high-quality headphones (Sennheiser HD 280, 64 ohm). Participants indicated their responses using a sensitive mouse (Logitech Corded Mouse M125 with high-definition optical tracking (1000-dpi)).

4.1.2 Results and discussion

Based on the RT distribution, responses with RTs longer than 5,000 ms were discarded from analyses for the identification test, which resulted in 302 responses (2.3%) being excluded for the affricate identification block, and 660 responses (5.1%) for the tone identification block. For the rating test, responses with RTs longer than 6,000 ms³⁹ were rejected from analyses, which resulted in 415 responses (3.7%) being excluded for the affricate rating block, and 622 responses (5.4%) for the tone rating block. In all ANOVAs described below, Mauchly's Test of Sphericity was used when the assumption of sphericity was violated. Degrees of freedom were corrected using Greenhouse-Geisser estimates of Sphericity.

³⁸A scale with an even number of steps was used to prevent participants from using the middle step excessively.

³⁹ When an RT 5,000 ms was used as a cut-off point for the rating test, the drop-out rate exceeded 6%, thus a longer RT cut-off, namely, 6,000 ms was used.

4.1.2.1 Adoptees vs. Dutch controls

4.1.2.1.1 Identification test

Figures 14 and Table 7, and Figure 15 and Table 8 show a distinct difference between the Cantonese adoptees and the Dutch controls: The affricates and tones produced by the Cantonese adoptees were identified more accurately than those produced by the non-adopted Dutch controls.

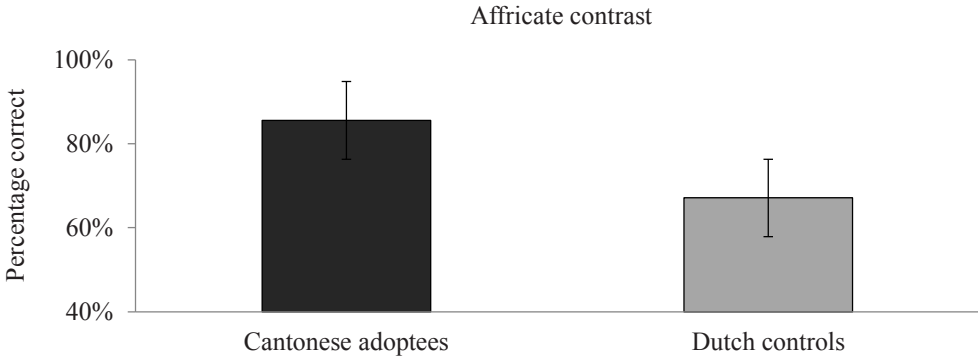


Figure 14. Percentage of correctly identified recordings for the Cantonese adoptees and the Dutch controls in the affricate identification block. Error bars represent standard errors.

Table 7

Percentage of correct for the Cantonese adoptees and the Dutch controls at all three test instances in the affricate identification block.

	Cantonese adoptees	Dutch controls
Pretest	81.8%	65.3%
Intermediate test	89.2%	67.6%
Posttest	85.9%	69.2%

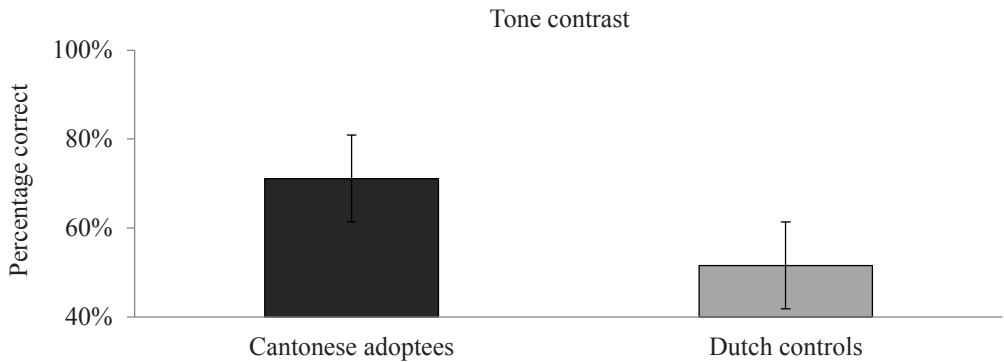


Figure 15. Percentage of correctly identified recordings for the Cantonese adoptees and the Dutch controls in the tone identification block. Error bars represent standard errors.

Table 8

Percentage of correct for the Cantonese adoptees and the Dutch controls at all three test instances in the affricate identification block.

	Cantonese adoptees	Dutch controls
Pretest	70.2%	51.7%
Intermediate test	72.3%	50.5%
Posttest	70.9%	52.5%

Proportions of correct responses were submitted to ANOVAs —by speakers (F_{IA}) and by native listeners (F_{IB})— with the independent variables⁴⁰ Group (Cantonese adoptees and Dutch controls), Test instance (Pretest, Intermediate test, and Posttest), and Target sound (Affricates: [ts] and [ts^h]; Tones: Tone 2 and Tone 5). As in Experiment 1 in Chapter 3 (section 3.1.2.1), the variable Parent Highest Education was included as a covariate in the analysis by participants.

Results confirmed the patterns described above. There was a significant Group difference for both the affricate and tone contrasts, showing that the Cantonese adoptees did significantly better than the Dutch controls in their production of the Cantonese affricate and tone contrasts across three test instances: Affricates: $F_{IA}(1, 41) = 22.352, p < .001, \eta^2_p = .353$, $F_{IB}(1, 23) = 1002.934, p < .001, \eta^2_p = .978$; Tones: $F_{IA}(1, 41) = 70.323, p < .001, \eta^2_p = .632$, $F_{IB}(1, 23) = 233.758, p < .001, \eta^2_p = .910$. Further, no effect of or interaction with Test instance, Target sound, or Parent Highest Education appeared in either block, which shows

⁴⁰ Group was both a between-subjects (F_{IA}) and a within-subjects (F_{IB}) variable, Test instance was a within-subjects variable, Target sound was a within-subjects variable, and Parent Highest Education was a covariate.

that the performance in neither group of children changed over time, and both target sounds of each contrast were pronounced similarly accurately. Importantly, F_{IA} ANOVAs similar to those in the main analysis above, but adding Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates (see also section 3.1.2.1) confirmed the results for both contrasts reported above⁴¹.

4.1.2.1.2 Rating test

The results of the rating test are very similar to those of the identification test. Figure 16 and Table 9 show a clear difference between the Cantonese adoptees and the Dutch controls for the affricate rating, but only for the unaspirated [ts], with the adoptees being rated higher than the controls. For the pronunciation of the aspirated [ts^h], the two groups of children were rated similarly.

Figure 17 and Table 10 also show a notable difference between the Cantonese adoptees and the Dutch controls for the tone rating: the adoptees received higher ratings than Dutch controls in the pronunciation of the Cantonese tone contrast. Additionally, Table 10 shows a difference in the production of Cantonese Tone 2 and Tone 5, with Tone 2 being rated higher than Tone 5.

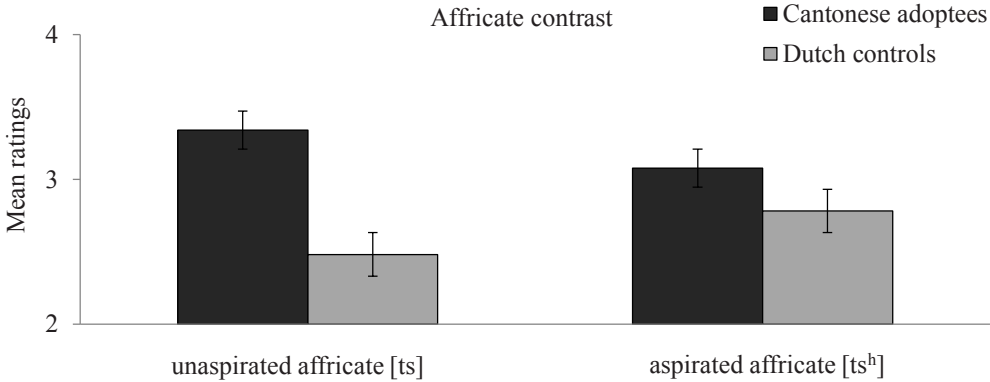


Figure 16. Mean ratings for the Cantonese adoptees and the Dutch controls in the affricate rating block. Ratings: 1: "completely different ", 2: "a little similar", 3: "very similar", 4: "identical". Error bars represent standard errors.

⁴¹ Additionally, for the tone identification, there was a significant interaction between Target sound and Age, $F_{IA}(1, 36) = 7.631, p < .01, \eta^2_p = .175$, and between Target sound and China Visits, $F_{IA}(1, 36) = 7.120, p < .05, \eta^2_p = .165$.

Table 9

Mean ratings for the Cantonese adoptees and the Dutch controls at all three test instances in the affricate rating block.

		Cantonese adoptees	Dutch controls
Pretest	unaspirated [ts]	3.2	2.4
	aspirated [ts^h]	3.0	2.8
Intermediate test	unaspirated [ts]	3.5	2.4
	aspirated [ts^h]	3.1	2.9
Posttest	unaspirated [ts]	3.4	2.6
	aspirated [ts^h]	3.1	2.6

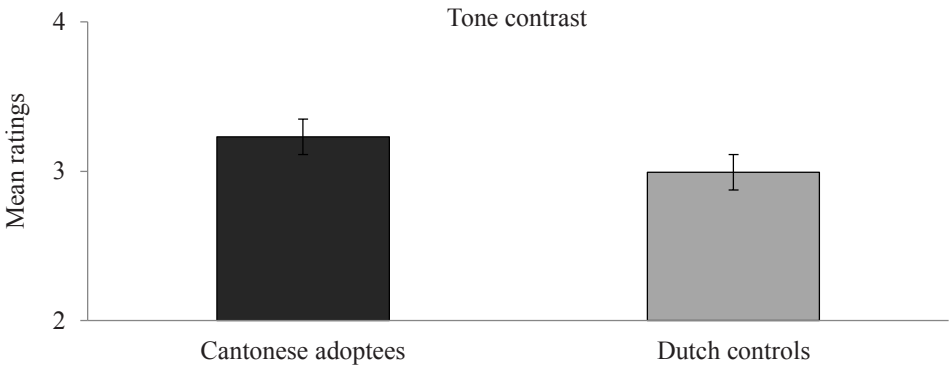


Figure 17. Mean ratings for the Cantonese adoptees and the Dutch controls in the tone rating block. Ratings: see Figure 16. Error bars represent standard errors.

Table 10

Mean ratings for the Cantonese adoptees and the Dutch controls at all three test instances in the tone rating block.

		Cantonese adoptees	Dutch controls
Pretest	Tone 2	3.5	3.2
	Tone 5	2.9	2.7
Intermediate test	Tone 2	3.5	3.3
	Tone 5	2.9	2.7
Posttest	Tone 2	3.5	3.3
	Tone 5	3.1	2.8

Similar ANOVAs to those described for the identification test were conducted for the rating test, but now with mean ratings as dependent variables. For the affricate rating, results showed a significant interaction between Group and Target sound, $F_{IA} (1, 41) = 5.865, p < .05, \eta^2_p = .125, F_{IB} (1, 20) = 131.967, p < .001, \eta^2_p = .868$. Follow-up analyses on the interaction revealed a significant Group difference for the unaspirated alveolar affricate [ts], with the Cantonese adoptees receiving significantly higher ratings than the Dutch controls, $F_{IA} (1, 42) = 27.446, p < .001, \eta^2_p = .395, F_{IB} (1, 20) = 499.252, p < .001, \eta^2_p = .961$, but no Group difference for the aspirated affricate [ts^h], $F_{IA} (1, 42) = 1.533, p > .05, \eta^2_p = .035, F_{IB} (1, 20) = 81.322, p < .001, \eta^2_p = .803$. There was no effect of Target sound either for the Cantonese adoptees, $F_{IA} (1, 21) = 3.126, p = .092, \eta^2_p = .130, F_{IB} (1, 20) = 19.865, p < .001, \eta^2_p = .498$, or for the Dutch controls, $F_{IA} (1, 21) = 2.924, p > .05, \eta^2_p = .122, F_{IB} (1, 20) = 18.492, p < .001, \eta^2_p = .480$. In addition, there was no effect of or interaction with either Test instance, or Parent Highest Education, indicating that neither group of children significantly improved their performance over time. Crucially, F_{IA} ANOVAs similar to those in the main analysis above, but adding Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates, again confirmed the main results reported above.

For the tone rating, there was a significant Group difference, showing that the Cantonese adoptees were rated significantly higher than the Dutch controls in their production of the tone contrast, $F_{IA} (1, 41) = 19.823, p < .001, \eta^2_p = .326, F_{IB} (1, 20) = 29.661, p < .001, \eta^2_p = .597$. Additionally, there was a significant effect of Target sound, showing that the production of Cantonese Tone 2 was rated significantly higher than Tone 5, $F_{IA} (1, 41) = 6.231, p < .05, \eta^2_p = .132, F_{IB} (1, 20) = 30.650, p < .001, \eta^2_p = .605$. Finally, there was no effect of or interaction with Test instance or Parent Highest Education. Again, neither group of children showed significant changes of their performance over time. Similar F_{IA} ANOVAs to those used in the main analysis above, but including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates, largely confirmed the main findings described above⁴².

In summary, the Cantonese adoptees were found to pronounce the Cantonese contrasts significantly better than the Dutch controls across all test instances (i.e., both before and after perceptual training). This finding was robust, with the one exception in the affricate

⁴² In addition, for the affricate rating, the interaction between Group and Target sound was no longer present, instead, a significant group difference was found ($F_{IA} (1, 36) = 11.839, p = .001, \eta^2_p = .247$), with the adoptees outperformed the controls. For the tone rating, there was a significant three-way interaction among Test instance, Target sound and Music, $F_{IA} (2, 72) = 4.488, p < .05, \eta^2_p = .111$.

rating block that the adoptees outperformed the controls only for the unaspirated affricate [ts]. However note that the group difference for the aspirated affricate [ts^h] was significant in the analysis across native listeners (F_{IB}); the fact that it was not in the analysis across speakers (F_{IA}) might be caused by large individual differences among the children. Furthermore, there were no significant changes in the production of the Cantonese contrasts over time either for the adoptees or for the controls. Possibly, the perceptual training program was too short to lead to any improvement (measurable with the present method). Crucially, the present findings show evidence for a long lasting effect of birth language experience on the production of segments and tones in the language.

4.1.2.2 Relationships among production and Age, AoA, and LoR

As in the perception study (in Chapter 3), the three 'time variables' Age, AoA, and LoR were also investigated in the present experiment with respect to their effects on the production of birth language contrasts for the Cantonese adoptees. AoA, LoR, and Age are linearly related, with a correlation between LoR and Age being inevitable, as described in section 3.1.2.3. Several measures of task performance were used: 1) Accuracy Overall, 2) at the Pretest, 3) at the Intermediate test, 4) at the Posttest, and 5) and 6) for each Target sound in each contrast (For the Cantonese alveolar affricates: [ts] and [ts^h]; For the Cantonese tones: Tone 2 and Tone 5).

Following the same procedures as in section 3.1.2.3, correlations were examined for both the identification and the rating scores, assessing the production performance measures versus Age, AoA and LoR (see Appendix F.2 for the statistical analyses). First, Kendall's Tau correlations were conducted between each of the six performance measures and each of the time variables. Results showed a significant correlation of Age (and LoR) with performance at the posttest for affricate identification. The older the Cantonese adoptees were at the time of testing (and hence the longer they had stayed in the Netherlands), the better identified was their production of the Cantonese affricate contrast. Age (and LoR) correlated significantly with several task performance measures for the affricate rating as well. The older the Cantonese adoptees were when tested (and the longer they had been in the Netherlands), the higher their production of the Cantonese affricate contrast was rated overall, at the pretest, and at the posttest⁴³. As discussed in Chapter 3, due to the inevitable linear relationship

⁴³ Note that LoR was only significantly correlated with the performance at the posttest for the affricate rating.

between Age and LoR, the unexpected effects of LoR are likely to be a reflection of the effects of Age. For the affricates, there were no significant correlations with AoA, either in the identification or in the rating test. For the tones, no significant correlation was found between any time variable and any performance measure, either for the identification or for the rating tests.

Next, partial correlations between each of the six performance measures, and Age and AoA, controlling for LoR, were carried out. Results showed no significant correlation either for the identification or for the rating test⁴⁴.

Finally, the 22 adoptees were divided into two sub-groups in order to disentangle the confounding of AoA and LoR, as in section 3.1.2.3 (see Appendix F.1.1). Thus, AoA and LoR were no longer correlated in either sub-group, although Age and LoR are of course still related. Again, Kendall's Tau correlations between each time variable and each measure of task performance in each sub-group were applied. For the affricate rating only, significant correlations were found, between Age (and LoR) and several task performances in the sub-group of 14 Cantonese adoptees. The older the Cantonese adoptees were when being tested (and the longer they had resided in the Netherlands), the higher their production of the Cantonese affricate contrast was rated overall, at the pretest, at the intermediate test, at the posttest, and particularly for the aspirated affricate [ts^h]⁴⁵. There were no other significant correlations.

Taken together, the present experiment showed no effect of AoA, but hints of evidence for an effect of Age (and LoR), particularly on the production of the Cantonese affricate contrast. However, note that no consistent findings of Age (and LoR) were found in both the identification and the rating tests, and –as discussed in Chapter 3– the sample size is small in comparison with the wide range of Age, AoA, and LoR. Results of Age (and LoR) should thus, again, be interpreted with caution.

4.1.2.3 Effects of Siblings and China Visits

As in Chapter 3 (section 3.1.2.4), two analyses were conducted to investigate the effects of 'Siblings' and 'China Visits' on the production of the Cantonese contrasts. Firstly, an F_{IA}

⁴⁴ Due to the linear relation ($AoA + LoR = Age$), the outcomes of the correlations with each performance measure were identical for Age and for AoA, after controlling for LoR.

⁴⁵ Note that LoR was not significantly correlated with the performance at the intermediate test, and for the aspirated affricate [ts^h], but it was significantly correlated for the unaspirated affricate [ts].

ANOVA was conducted similar to that in the main analyses in section 4.1.2.1, but including Siblings and China Visits as covariates. Results confirmed the main findings in both identification and rating tests. Interestingly, there was a key effect of Siblings ($F_{IA}(1, 40) = 5.405, p < .05, \eta^2_p = .119$) for the affricate rating, showing that the children with adopted Cantonese siblings were rated to have significantly better production of the Cantonese affricate contrast compared to those without adopted siblings. Further, there was a significant interaction between Test instance and China Visits ($F_{IA}(2, 80) = 5.086, p < .01, \eta^2_p = .113$) for the tone rating. Follow-up analyses on the interaction revealed a significant effect of Test instance for the children who had visited the Cantonese areas in China, showing that the production of the Cantonese tone contrast was rated significantly higher in the posttest in comparison with the intermediate test, and the pretest. Finally, there was a three-way interaction among Test instance, Target sound, and Siblings ($F_{IA}(2, 80) = 4.396, p < .05, \eta^2_p = .099$) for the tone identification.

Second, an F_{IA} ANOVA was applied similar to that described above, but adding Siblings and China Visits as independent variables and leaving out all covariates. The results are consistent with the findings of the main analyses in section 4.1.2.1. Importantly, the results also confirmed those described in the paragraph above. Specifically, there was a significant effect of Siblings ($F_{IA}(1, 37) = 5.584, p < .05, \eta^2_p = .131$) for the affricate rating, with children who have adopted Cantonese siblings receiving higher ratings in their production of the Cantonese affricate contrast than those without adopted Cantonese siblings. There was also a significant interaction between Test instance and China Visits ($F_{IA}(2, 74) = 3.465, p < .05, \eta^2_p = .086$) for the tone rating, with children who had visited the Cantonese areas in China receiving higher ratings for their production of the Cantonese tone contrast at the posttest, compared to the pretest and intermediate test ($F_{IA}(2, 22) = 7.912, p < .01, \eta^2_p = .418$). Further, there was a significant three-way interaction among Test instance, Target sound, and Siblings ($F_{IA}(2, 74) = 4.848, p < .05, \eta^2_p = .116$) for the tone identification. Additionally, the present analysis showed a significant effect of China Visits ($F_{IA}(1, 37) = 4.252, p < .05, \eta^2_p = .103$) for the affricate rating, with children who had visited the Cantonese areas in China receiving higher ratings than those who had not visited China on the production of the Cantonese affricate contrast. There was another significant three-way interaction among Test instance, Target sound, and Siblings ($F_{IA}(2, 74) = 4.425, p < .05, \eta^2_p = .085$) for the tone rating. Taken together, the present evidence suggests, consistent with our

expectations, that having adopted siblings and having visited China each have positive effects on the production of the Cantonese contrasts.

4.1.2.4 Correlations between perception and production

It was investigated whether speech perception and production of the birth language contrasts were correlated, for both the adoptees and the controls. To this end, six performance measures, the same as those in section 3.1.2.3, were used separately for the perception experiment and the production assessment tests (i.e., identification and rating tests). Kendall's Tau correlations were applied for each performance measure between the perception experiment and each production assessment test. Results showed no significant correlations either for the adoptees or for the controls, with one exception: the Cantonese adoptees' perception of the tone contrast in the intermediate test was positively correlated with their production of the tone contrast in the intermediate test in the rating test. The present data show no clear relationship between the perception and the production of the Cantonese contrasts.

4.1.2.5 Summary

Consistent results from both the identification and the rating tests showed a robust advantage (with just one exception) in the Cantonese adoptees over the Dutch controls in production of the Cantonese contrasts across all test instances (before, between and after perceptual training blocks). Further, whereas no effect of AoA was found in either test, small effects of Age (and LoR) were observed, particularly on the production of the affricates. In addition, there is a trend in the data for having adopted siblings, and having visited the Cantonese areas in China, to be related to better production of the Cantonese contrasts. Finally, the production and the perception of the Cantonese contrasts appeared not to be related.

4.2 Experiment 2: Mandarin production

This experiment tested the production of the Mandarin affricate and tone contrasts. As in Experiment 1, two types of contrasts were tested in separate parts (see Table 3 in Chapter 2). The present experiment aims to replicate and extend the findings of Experiment 1.

4.2.1 Method

4.2.1.1 Participants

Participants were the 26 Mandarin adoptees and the 26 Dutch controls who were also tested in the Mandarin perception experiment in Chapter 3.

4.2.1.2 Materials

4.2.1.2.1 Speech materials

As in Experiment 1, production testing involved two high-frequency⁴⁶ minimal pairs of Mandarin words for each contrast, chosen from the 16 (per contrast) that were tested in perception (see Table 11). Each elicitation stimulus used two identical tokens (from the same speaker), thereby resulting in eight tokens per contrast. Again, only the two speakers who were chosen for the baby animals in the video game were used (see section 2.2.2).

Table 11

Test stimuli of each experimental contrast in Experiment 2 with Chinese characters.

Retroflex affricate contrast		Tone contrast	
unaspirated	aspirated	Tone 2 (High-Rising)	Tone 3 (Low-Rising)
[atʂau] "阿招"	[atʂ ^h au] "阿超"	[amei] "阿梅"	[amei] "阿美"
[atʂuan] "阿专"	[atʂ ^h uan] "阿川"	[at ^h ien] "阿甜"	[at ^h ien] "阿舔"

Note. The first syllable [a] "阿" in both contrasts, and the second syllable in the affricate contrast are in tone 1, while the second syllable in the tone contrast is either in tone 2 or tone 3.

⁴⁶ Two native speakers of Mandarin (*Age* = 32 and 30 years old) chose these two minimal pairs as being, according to their subjective assessment, the most frequent items from among the 16, for the children in the age range under study.

4.2.1.2.2 Video materials

The same video materials were used as in Experiment 1.

4.2.1.3 Design and procedure

The same design and procedure as those in Experiment 1 were used. In total, however, the assessment experiment contained 624 stimuli⁴⁷ per contrast (4 words * 3 test instances * 52 participants), because the number of children was larger in Experiment 2 than in Experiment 1.

4.2.1.4 Assessment of children's productions

4.2.1.4.1 Participants

Forty-four native Mandarin Chinese listeners, who were born and had lived in the Mandarin Chinese speaking areas in the north of China, participated in the assessment experiment. At the time of testing, all participants were all college students⁴⁸ studying in Beijing. They were also randomly assigned to two groups for the two different assessment tasks (described in section 4.1.1.4.2) (see Table 12). Each of them also received a small amount of money for participating.

Table 12

Descriptive information of the native Mandarin participants in the assessment experiment.

	Number	Mean age	Gender		Education	
Identification	22	23;5 years	11 F	11 M	10 BA	12 MA
Rating	22	23;1 years	11 F	11 M	10 BA	12 MA

Note. F": Female, and "M": Male. Education: The highest education on-going at the time of testing. "BA": Bachelor program, and "MA": Master program.

⁴⁷ Note that only the first token of the target word was used in the assessment experiment, as in Experiment 1.

⁴⁸ Participants were recruited from seven different universities in Beijing, namely Peking University, Tsinghua University, Beijing Normal University, Renmin University of China, Beijing Institute of Technology, Beijing Foreign Studies University, and University of Science and Technology.

4.2.1.4.2 Materials

As in Experiment 1, the 624 stimuli per contrast (see section 4.2.1.3) collected from the Mandarin adoptees and the Dutch controls were used in the assessment experiment. The same 2AFC identification and rating tasks were used as in Experiment 1.

4.2.1.4.3 Procedure

The procedure was as in Experiment 1. Each test block, however, here lasted 28 - 30 minutes (due to the larger number of stimuli).

4.2.2 Results and discussion

Responses with RTs longer than 5,000 ms were considered as outliers, and removed from the main analyses for both the identification and the rating tests, which led to 243 responses (1.8%) being excluded from the affricate identification block, 399 responses (2.9%) from the tone identification block, 410 responses (3%) from the affricate rating block, and 499 responses (3.6%) from the tone rating block. For all ANOVAs conducted below, Mauchly's Test of Sphericity showed that the assumption of sphericity had not been violated.

4.2.2.1 Adoptees vs. Dutch controls

4.2.2.1.1 Identification test

For the affricates, Figure 18 and Table 13 show a major difference between the two groups of children, with higher scores for the Mandarin adoptees than the Dutch controls. In addition, Table 13 shows that the production of the Mandarin aspirated affricate [tʂ^h] was scored higher than that of the unaspirated affricate [tʂ].

For the tones, Figure 19 and Table 14 show a less general but yet a considerable difference between the two groups of children: the adoptees were scored higher than the controls, but only in the production of Mandarin Tone 3. The two tones also differed significantly: Tone 2 was pronounced much more accurately, by all participants, than Tone 3.

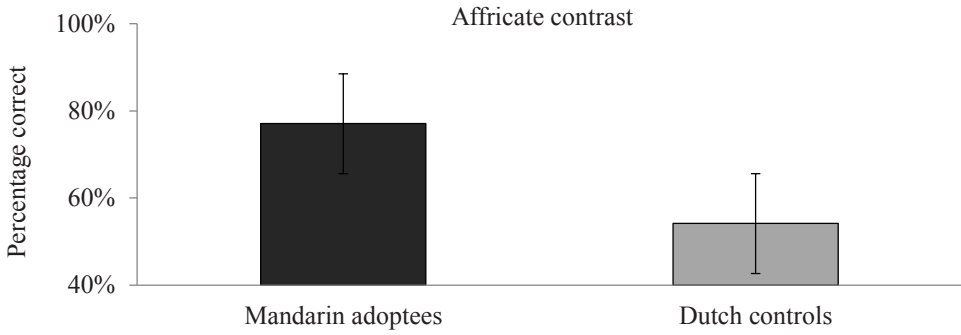


Figure 18. Percentage of correctly identified recordings for the Mandarin adoptees and the Dutch controls in the affricate identification block. Error bars represent standard errors.

Table 13

Percentage of correct for the Mandarin adoptees and the Dutch controls at all three test instances in the affricate identification block.

		Mandarin adoptees	Dutch controls
Pretest	unaspirated [tʂ]	62.5%	39.6%
	aspirated [tʂ^h]	88.1%	71.3%
Intermediate test	unaspirated [tʂ]	67.3%	41.7%
	aspirated [tʂ^h]	89.2%	68.1%
Posttest	unaspirated [tʂ]	68.8%	43.8%
	aspirated [tʂ^h]	86.6%	60.5%

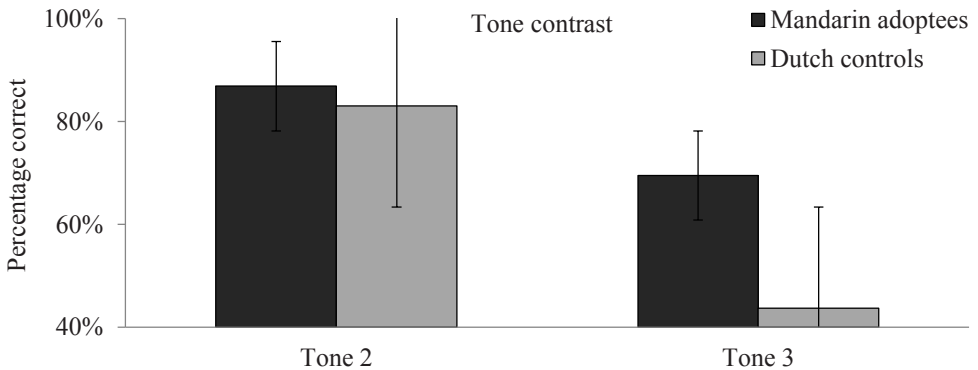


Figure 19. Percentage of correctly identified recordings for the Mandarin adoptees and the Dutch controls in the tone identification block. Error bars represent standard errors.

Table 14

Percentage of correct for the Mandarin adoptees and the Dutch controls at all three test instances in the tone identification block.

		Mandarin adoptees	Dutch controls
Pretest	Tone 2	87.4%	81.4%
	Tone 3	68.8%	43.9%
Intermediate test	Tone 2	85.4%	84.3%
	Tone 3	64.0%	41.4%
posttest	Tone 2	87.8%	83.3%
	Tone 3	75.7%	45.8%

All the patterns discussed above are corroborated by the statistical analyses. The same ANOVAs were used, as in Experiment 1, but without covariates (because all control variables were well-matched). Confirming Figure 19, there was a significant Group difference for the affricate identification, $F_{1A}(1, 50) = 69.322, p < .001, \eta^2_p = .581, F_{1B}(1, 21) = 1900.316, p < .001, \eta^2_p = .989$, showing that the Mandarin adoptees outperformed the Dutch controls in their production of the Mandarin affricate contrast across all test instances. Further, there was a significant effect of Target sound, $F_{1A}(1, 50) = 19.986, p < .001, \eta^2_p = .286, F_{1B}(1, 21) = 58.438, p < .001, \eta^2_p = .736$, showing that the Mandarin aspirated affricate [tʂ^h] was identified more accurately than the unaspirated affricate [tʂ]. No effect of or interaction with Test instance was found, suggesting that the performance of neither group of children changed significantly over time. Similar F_{1A} ANOVAs to those used in the main analysis above, but including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates, largely confirmed the main results described above⁴⁹.

For the tone identification, results show a significant interaction between Group and Target sound, $F_{1A}(1, 50) = 8.863, p = .004, \eta^2_p = .151, F_{1B}(1, 21) = 253.005, p < .001, \eta^2_p = .923$. Follow-up analyses on the interaction showed a significant Group difference for Mandarin Tone 3, with the adoptees being scored significantly better than the controls, $F_{1A}(1, 50) = 25.897, p < .001, \eta^2_p = .341, F_{1B}(1, 21) = 394.297, p < .001, \eta^2_p = .949$, but no Group difference for Mandarin Tone 2, $F_{1A}(1, 50) = 1.321, p > .05, \eta^2_p = .026, F_{1B}(1, 21) = 27.099$,

⁴⁹ In addition, there was a significant interaction between Target sound and China Visits, $F_{1A}(1, 44) = 8.601, p < .01, \eta^2_p = .164$, and between Target sound and Siblings, $F_{1A}(1, 44) = 4.779, p < .05, \eta^2_p = .098$. There was also a significant effect of Age, $F_{1A}(1, 44) = 7.215, p < .05, \eta^2_p = .141$, of Gender, $F_{1A}(1, 44) = 6.030, p < .05, \eta^2_p = .121$, and of Music, $F_{1A}(1, 44) = 4.823, p < .05, \eta^2_p = .099$.

$p < .001$, $\eta_p^2 = .563$. Additionally, there was a significant effect of Target sound, with Mandarin Tone 2 being produced significantly better than Mandarin Tone 3 by both the Mandarin adoptees: $F_{IA} (1, 25) = 21.326$, $p < .001$, $\eta_p^2 = .460$, $F_{IB} (1, 21) = 20.376$, $p < .001$, $\eta_p^2 = .492$, and the Dutch controls: $F_{IA} (1, 25) = 38.676$, $p < .001$, $\eta_p^2 = .607$, $F_{IB} (1, 21) = 104.418$, $p < .001$, $\eta_p^2 = .833$. Finally, there was no effect of or interaction with Test instance, showing that neither the adoptees nor the controls improved their performance over time. Again, similar F_{IA} ANOVAs to those conducted in the main analyses above, but including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates, basically confirmed the main findings reported above⁵⁰.

4.2.2.1.2 Rating test

The results of the rating test strongly resemble those of the identification test. As shown in Figure 20 and Table 15, there was a big difference between the adoptees and the controls for the affricate rating, showing that the adoptees received higher ratings than the Dutch controls. As also shown in Table 15, there was a difference in the production of the two affricate sounds, with the aspirated [tʂ^h] being rated higher than the unaspirated [tʂ].

As shown in Figure 21 and Table 16, there was also a clear difference between the adoptees and the controls for the tone rating, showing that the adoptees were rated higher than the controls. Table 16 also shows that the production of two tones was rated differently, with Mandarin Tone 2 higher than Tone 3.

⁵⁰ The only additional effect was a significant interaction between Test instance and Age, $F_{IA} (2, 88) = 3.988$, $p < .05$, $\eta_p^2 = .083$.

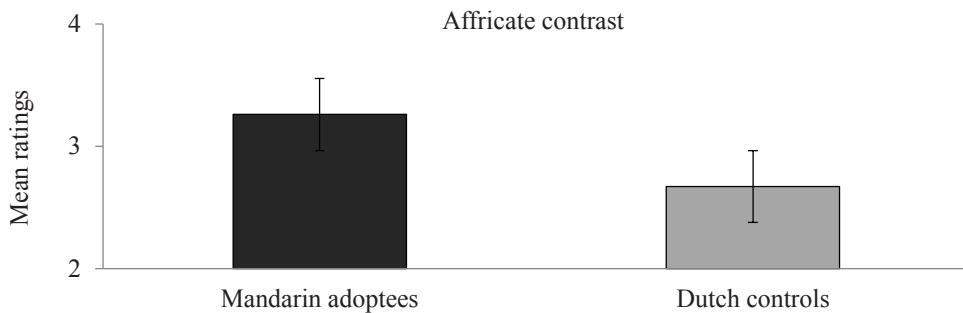


Figure 20. Mean ratings for the Mandarin adoptees and the Dutch controls in the affricate rating block. Ratings: see Figure 16. Error bars represent standard errors.

Table 15

Mean ratings for the Mandarin adoptees and the Dutch controls at all three test instances in the affricate rating block.

		Mandarin adoptees	Dutch controls
Pretest	unaspirated [tʂ]	3.1	2.5
	aspirated [tʂ^h]	3.3	2.9
Intermediate test	unaspirated [tʂ]	3.2	2.5
	aspirated [tʂ^h]	3.3	2.8
Posttest	unaspirated [tʂ]	3.3	2.6
	aspirated [tʂ^h]	3.4	2.7

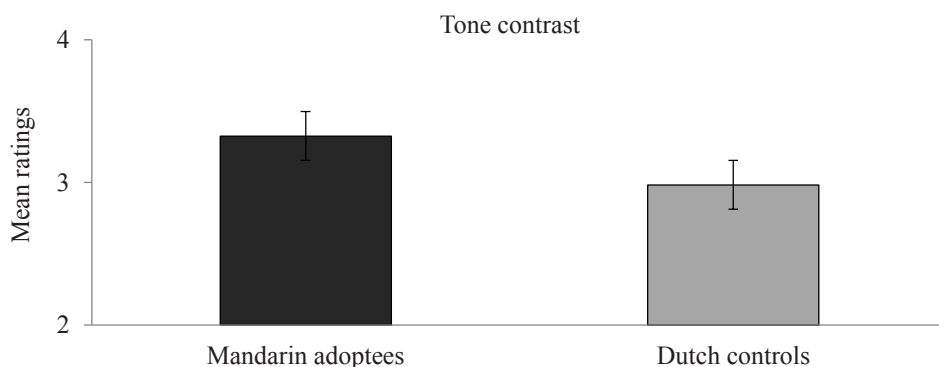


Figure 21. Mean ratings for the Mandarin adoptees and the Dutch controls in the tone rating block. Ratings: see Figure 16. Error bars represent standard errors.

Table 16

Mean ratings for the Mandarin adoptees and the Dutch controls at all three test instances in the tone rating block.

		Mandarin adoptees	Dutch controls
		Pretest	Tone 2
	Tone 3	3.0	2.4
Intermediate test	Tone 2	3.7	3.6
	Tone 3	2.9	2.4
Posttest	Tone 2	3.7	3.6
	Tone 3	3.1	2.4

ANOVAs were carried out as in section 4.1.2.1.2, but without covariates; these confirmed all patterns described above. For the affricate rating, there was a significant Group difference, with the Mandarin adoptees’ production being rated significantly higher than that of the Dutch controls, $F_{IA}(1, 50) = 71.129, p < .001, \eta^2_p = .587, F_{IB}(1, 21) = 281.689, p < .001, \eta^2_p = .931$. In addition, there was a significant effect of Target sound, with the aspirated affricate [tʂ^h] being produced significantly more acceptably than the unaspirated affricate [tʂ], $F_{IA}(1, 50) = 4.337, p < .05, \eta^2_p = .080, F_{IB}(1, 21) = 12.542, p = .002, \eta^2_p = .374$. There was no effect of or interaction with Test instance, which indicates that neither the adoptees nor the controls showed any change in performance across test instances. Essentially similar F_{IA} ANOVAs as for the main analyses above, but including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates, largely confirmed the main results presented above⁵¹.

The tone ratings displayed a significant interaction between Group and Target sound, $F_{IA}(1, 50) = 13.280, p = .001, \eta^2_p = .210, F_{IB}(1, 21) = 99.884, p < .001, \eta^2_p = .826$. Follow-up analyses on the interaction showed that, first, there was a significant Group difference - with the adoptees being scored significantly higher than the controls - for both Mandarin Tone 2, $F_{IA}(1, 50) = 6.904, p < .05, \eta^2_p = .121, F_{IB}(1, 21) = 64.544, p < .001, \eta^2_p = .755$, and Tone 3, $F_{IA}(1, 50) = 26.108, p < .001, \eta^2_p = .343, F_{IB}(1, 21) = 154.239, p < .001, \eta^2_p = .880$. Second, there was a significant effect of Target sound - with the production of Mandarin Tone 2 being rated significantly better than that of Tone 3 - for both the Mandarin adoptees:

⁵¹ In addition, for the affricate rating, the main effect of Target sound was now marginally significant, $F_{IA}(1, 44) = 3.961, p = .053, \eta^2_p = .083$. Moreover, there was a significant interaction between Target sound and Siblings, $F_{IA}(1, 44) = 5.011, p < .05, \eta^2_p = .102$, between Target sound and China Visits, $F_{IA}(1, 44) = 6.535, p < .05, \eta^2_p = .129$.

$F_{IA}(1, 25) = 83.385, p < .001, \eta^2_p = .769, F_{IB}(1, 21) = 109.433, p < .001, \eta^2_p = .839$, and the Dutch controls: $F_{IA}(1, 25) = 117.295, p < .001, \eta^2_p = .824, F_{IB}(1, 21) = 152.229, p < .001, \eta^2_p = .879$. Again the main findings above are confirmed when similar F_{IA} ANOVAs to those in the analyses above, but including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates.

The results of Experiment 2 are in line with Experiment 1, showing that the Mandarin adoptees had better production of their birth language contrasts than the Dutch controls in both the identification and the rating tests across all test instances (i.e., before, between, and after perceptual training). One exception to this general pattern was found in the tone identification block, where the adoptees outperformed the controls only for Mandarin Tone 3, but not for Mandarin Tone 2. Note that the group difference for Mandarin Tone 2 was significant in the analysis across native listeners (F_{IB}); but not in the analysis across speakers (F_{IA}), which might be due to the great individual differences among the children, as discussed for Experiment 1. Again, no significant improvements over time were observed in either group of children, presumably due to the short training program. Further, whereas no clear difference had appeared in the production of the target sounds in Experiment 1, in the present experiment, both the identification and the rating tests consistently showed a distinct difference in the production of the two target sounds in each contrast. This will be discussed below. Together with Experiment 1, the present findings confirmed a long lasting benefit of early birth language experience for production of the sounds of the language.

4.2.2.2 Relationships among production and Age, AoA, and LoR

As for Experiment 1, correlational analyses were conducted to investigate the effects of the three 'time variables' (Age, AoA, and LoR) on production (see Appendix F2 for the statistical analyses). Recall that Age, AoA, and LoR are linearly related, and the correlation between Age and LoR is non-disentangled. The same six performance measures as in Experiment 1 were used: 1) Accuracy Overall, 2) at the Pretest, 3) at the Intermediate test, 4) at the Posttest, and 5) and 6) for each Target sound in each contrast (For the Mandarin retroflex affricates: [tʂ] and [tʂ^h]; For the Mandarin tones: Tone 2 and Tone 3).

First, Kendall's Tau correlational analyses between each of the six performance measures and each of the three time variables showed several significant positive and negative correlations. *Positive correlations*: Older adoptees at the time of testing (who had

been in the Netherlands longer) were rated to have better production of the Mandarin tone contrast at the posttest in the rating test. The adoptees who were adopted at an older age were identified to produce Mandarin Tone 2 better in the identification test. *Negative correlations:* In the identification test, the older the adoptees were when tested, the less accurately their production of the Mandarin affricate contrast was assessed at the pretest. The adoptees with longer residence in the Netherlands were scored lower in their production of the Mandarin affricate contrast overall and at the pretest, and particularly for the production of the unaspirated affricate [tʂ]. Further, adoptees who had been in the Netherlands longer were evaluated to produce Mandarin tone 2, and the tone contrast less accurately at the pretest.

Next, partial correlations between each time variable and each task performance, controlling LoR, were conducted, and showed no significant correlations in either the identification or the rating test⁵².

Finally, the 26 adoptees⁵³ were divided into two sub-groups, as before teasing apart the correlation between AoA and LoR⁵⁴, (see Table F.1.3 in Appendix F1). Again, Kendall's Tau correlations among the production task performance and Age, AoA, and LoR in each sub-group for both the identification and the rating tests were carried out. Only in the case of the sub-group of 19 adoptees for the identification test, Age (and LoR) was *negatively* correlated with performance at the pretest for the affricate contrast, and performance for Tone 2. No other correlations were found.

In line with Experiment 1, no clear evidence of an effect of AoA was found. In contrast to Experiment 1 which showed a tentative positive effect of Age (and LoR) on the overall production of the Cantonese affricate contrast, Experiment 2 mainly showed negative effects of Age and LoR on the production of the Mandarin contrasts, particularly before training. The negative effects of Age and LoR suggest that the adoptees who were older at the time of testing, and those who had been cut off from their birth language longer produced their birth language contrasts less accurately, before re-exposure.

⁵² Due to the linear relation ($AoA + LoR = Age$), the outcomes of the correlations with each performance measure were identical for Age and for AoA, after controlling for LoR.

⁵³ As described in section 3.2.2.3, one adoptee was excluded from the re-grouping procedure as an outlier because he was much older than the other children at the time of adoption.

⁵⁴ Note that AoA and LoR were no longer correlated in either sub-group, but Age and LoR were still significantly correlated.

4.2.2.3 *Effects of Siblings and China Visits*

As in Experiment 1, the effects of 'Siblings' and 'China Visits' on the production of the Mandarin contrasts were assessed in the following two analyses. First, an F_{IA} ANOVA similar to that in section 4.2.2.1, but including Siblings and China Visits as covariates, confirmed the findings described in section 4.2.2.1, but it showed no effect of or interaction with either Siblings or China Visits in either the identification or the rating test.

Second, a similar F_{IA} ANOVA to that described above, but using Siblings and China Visits as independent variables rather than covariates, largely confirmed the main results in section 4.2.2.1⁵⁵. Additionally, a significant interaction between Group and China Visits was found in the tone rating task, $F_{IA}(1, 45) = 4.078, p < .05, \eta^2_p = .083$ ⁵⁶. Follow-up analyses on the interaction revealed a significant group difference for the children who had never visited the Mandarin areas of China, with the adoptees outperforming the Dutch controls, $F_{IA}(1, 40) = 18.933, p < .001, \eta^2_p = .321$. For the children who had visited the Mandarin areas in China, no group difference was found, $F_{IA}(1, 5) = .124, p > .05, \eta^2_p = .024$.

4.2.2.4 *Correlations between perception and production*

As in Experiment 1, the relationship between the perception and the production of the Mandarin contrasts was assessed. The same six task performance measures as those in section 3.2.2.3 were used for both the perception experiment and the production assessment tests. Largely confirming Experiment 1, Kendall's Tau correlational analysis showed no significant correlation, for either the adoptees or the controls in Experiment 2, with one exception: in the affricate rating test, the Dutch controls' production of the Mandarin affricate contrast at the pretest positively correlated with their perception of the affricate contrast at the pretest.

4.2.2.5 *Summary*

Overall, Experiment 2 confirmed the pattern shown in Experiment 1: the adoptees were better than the Dutch controls in producing their birth language contrasts, across all test instances (before, between, and after perceptual training), with the one exception for Mandarin Tone 2 in the identification task. Further, no significant improvement was observed across the

⁵⁵ Note that in the affricate rating test, the effect of Target sound was no longer present. In both the tone identification and rating tests, the interaction between Group and Target sound was also gone, but a three-way interaction among Group, Target sound, and China Visits was found.

⁵⁶ A significant interaction between Sibling and China Visits were found.

training blocks in either group of children, as in Experiment 1. In contrast to Experiment 1, however, Experiment 2 showed a distinct difference in the production of the two target sounds in each contrast, which is further discussed below. As for the three 'time variables', Experiment 2, like Experiment 1, showed no clear effect of AoA (again with Mandarin Tone 2 being an exception). However, contrary to Experiment 1 in which small positive effects of Age and LoR occurred in the production of the affricate contrast in particular, Experiment 2 generally showed negative effects of Age and LoR on the production of the Mandarin contrasts, particularly before training. Again counter to Experiment 1, no clear evidence for the effects of Siblings and China Visits were found in Experiment 2. Finally, in line with Experiment 1, no clear evidence supports a relation between the perception and the production of the Mandarin contrasts (with one exception).

4.3 General discussion

4.3.1 Long-term benefits of early phonological experience

The present study found converging evidence of better production of Chinese contrasts in the Chinese adoptees compared to the Dutch controls in both Experiments 1 and 2, across all test instances. These findings are in keeping with the study on *adult* Korean adoptees (Choi, 2014) and the studies on *adult* heritage language learners (e.g., Au et al., 2002, 2008; Knightly et al., 2003; Oh et al., 2003), in which participants showed an advantage in the production of the phonemes from their childhood languages, although only *after* relearning. The most striking aspect of the present results therefore is that they extend previous findings by revealing advantages even *before* perceptual training in the young adopted Chinese children. This difference between the present study and those previous studies points to the importance of the time elapsed after adoption for birth language maintenance. Together with the results of the perception study in Chapter 3, the present findings corroborate that early birth language experience has long-term benefits; moreover, the present study shows that benefits appear not only in perception but also in production of the language.

The high accuracy of the Chinese adoptees' performance across test instances is especially remarkable, in comparison with the results of the Dutch controls. The adoptees outperformed the controls for both the affricate and tone contrasts in both the identification and the rating tests (with two exceptions for the production of the Cantonese aspirated

alveolar affricate [ts^h] in the rating test, and for Mandarin Tone 2 in the identification test, where differences were insignificant. The controls never outperformed the adoptees.). As discussed earlier, the group difference in those two test blocks was significant across native listeners (F_{IB}), but not significant across speakers (F_{IA}), which can be ascribed to the large individual differences among the children.

In contrast to previous studies which showed improved production as a result of perceptual training, the present study found that neither the adoptees nor the controls significantly improved their production through the perceptual training sessions. Previous studies suggest that phonological knowledge acquired through perceptual training can be transferred to production. For instance, Japanese adults were reported not only to have improved their perception but also their production of the English [r] and [l] contrast through intensive perceptual training (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997). Similarly, American English learners also showed improvement in production of Mandarin tone contrasts after perceptual training (Wang, Jongman, & Sereno, 2003). Note however, that there are methodological differences between the previous research with adults and our study, e.g., that the perceptual training in the present study was rather short because of the young age of the participants.

Interestingly, there was also evidence that one target sound was produced better than the other, particularly within each *Mandarin* contrast, which seems to suggest that the two target sounds in each contrast were not equally easy to produce. The Mandarin aspirated retroflex affricate [tʂ^h] was produced better than the unaspirated [tʂ], in both the identification and the rating tests. This result is rather different from the longitudinal case study by Zhu (2002), in which the unaspirated [tʂ] appeared earlier than the aspirated [tʂ^h] in native children's speech; thus, the pattern in the present study might not point to the aspirated affricate being easier to articulate than the unaspirated one, but might result from a relation with Dutch sounds (see Flege, 1995, Speech Learning Model). Furthermore, Mandarin Tone 2 was produced better than Tone 3, in both the identification and the rating tests. This result is in line with the longitudinal case study by Zhu (2002), in which Mandarin Tone 2 was reported to be easier to produce than Tone 3. As for the Cantonese contrasts, only in the rating test, Cantonese Tone 2 was produced better than Tone 5, which is consistent with the findings for native Cantonese children in (So & Dodd, 1995; Tse, 1978).

4.3.2 Effects of AoA and Age (and LoR)

Within the groups of adoptees, opposite effects of Age and LoR were found in the two experiments. In experiment 1, the older the Cantonese adoptees were when tested and the longer they had lived in the Netherlands, the *more accurate* their productions of the Cantonese affricate contrast were evaluated, particularly *after* perceptual training. This might be because older adoptees might have learned more because of their cognitive maturity, and/or because the task might have been easier for them with respect to various non-linguistic task demands (e.g., sustained attention, pressing the button, following instructions, etc.). In Experiment 2, the older the Mandarin adoptees were at the time of testing and the longer they had been in the Netherlands, the *less accurate* their production of the Mandarin contrasts were assessed, particularly *before* perceptual training. This suggests that older adoptees, who had been separated from their birth language longer, became less fluent in producing their birth language contrasts before relearning. For the Mandarin tone rating only, Age and LoR were positively correlated with production of the Mandarin tone contrast in the posttest. AoA had no consistent effect, with one exception for Mandarin Tone 2 in the identification test, in which adoptees who were adopted at an older age produced Mandarin Tone 2 better. The present study thus shows highly variable effects of Age (and LoR) and/or AoA. As discussed in Chapter 3, the sample size in the present study is rather small for correlational analyses, which might explain this variability.

Within the groups of Dutch controls, small effects of Age were found in both experiments. In experiment 1, the older the controls were, the better they produced the Cantonese tone contrast particularly after perceptual training, and the better they produced Cantonese Tone 2. In experiment 2, the older the controls were, the better they produced the Mandarin unaspirated retroflex affricate [tʂ] (only in the rating block). These results seem to confirm that older control children were more task ripe (as discussed for the older adoptees), leading to better performance after learning. However, note that the results are not replicated in the identification and the rating tests of the two experiments. Furthermore, as discussed above, the size of the control group in each experiment was small for correlational analyses. To develop a better understanding of the effect of Age on the production of the Chinese contrasts, further studies with a bigger sample size will be needed.

4.3.3 Effects of Siblings and China Visits

The effects of Siblings and China Visits were different in Experiments 1 and 2. In Experiment 1, children with adopted Cantonese siblings produced the affricate contrast better than those without adopted siblings (only for the affricate rating task). In Experiment 2, no effect of Siblings was found in either identification or rating.

As for China Visits in Experiment 1, children who had visited the Cantonese areas in China showed better production of the affricate contrast than those who had never visited China (only for affricate rating⁵⁷). In addition, children who had visited the Cantonese areas in China, but not those children who had not, showed improvement of production of the *tone* contrast, with better production in the posttest as compared to the pretest and the intermediate test, particularly in the rating task. As discussed in Chapter 3, two possible reasons accounted for the positive effect of China Visits: First, the children who had visited the Cantonese areas in China could have overheard Cantonese sounds (re-exposure for the adoptees), which may have helped them produce the contrast. Second, the children who had visited the Cantonese areas in China might be more motivated in the experiment. As for China Visits in Experiment 2, the picture is rather unclear. For tone rating only⁵⁸, China Visits was found to interact with Group. Among the children who had never visited the Mandarin areas in China, the adoptees produced the Mandarin tone contrast better than the controls. This result suggests that the early birth language experience of those adoptees facilitated production of the tone contrast in their birth language, such that they outperformed the controls without prior experience with tones. Among the children who had visited the Mandarin areas in China, the adoptees and the controls performed similarly. This seems to suggest that the exposure to the Mandarin tone contrast during the visits in China - although limited to a few days or weeks - benefited the Dutch controls in their production of the Mandarin tone contrast, and thus helped them to catch up with the adoptees. However, note that no consistent effect of China Visits was found either for the Mandarin tone identification, or for the affricate contrast, such that caution must be applied when interpreting the result of China Visits in Experiment 2.

⁵⁷ The effect of China Visits was found only in the F_1 analysis with Siblings and China Visits as independent variables.

⁵⁸ It is also the F_1 analysis with Siblings and China Visits as independent variables.

4.3.4 No correlation between perception and production

There was no clear evidence for a connection between the perception and the production of the Chinese contrasts (with two exceptions) in either experiment. The difference between the outcomes of the perception and the production studies, particularly at the pretest and the intermediate test, is striking. Whereas in the production study, both the Cantonese and the Mandarin adoptees produced their birth language contrasts significantly better than the Dutch controls across all test instances, in the perception study the adoptees outperformed the controls only after completing all perceptual training blocks.

To conclude, the present study has demonstrated, for the first time, a clear advantage in the production of birth language contrasts before and after re-exposure to the language by a large number of young adopted children. The findings from the present study make several contributions to the current literature. First, previous evidence of the continuing influence of early birth language experience is confirmed; not only in the domain of speech perception but also in the domain of speech production can such influence be seen. Second, our understanding of the effect of time elapsed post adoption is deepened. Third, the effects demonstrated here of age, length of residence in the country of adoption, and age of adoption, even though variable, should prove valuable in future research. Fourth, exposure to birth-language sounds, either from siblings who were also adopted, or from visits to the country of birth, seems to be beneficial for producing the sounds of the birth language. And finally, this study provides the first empirical evidence on the relationship of perception and production of birth language contrasts in international adoptees.

Chapter 5: Residual memory of birth language vocabulary

The research reported in this chapter assesses the adopted Chinese children's residual memories of the vocabulary of their birth language, and again their test performance is compared to that of both non-adopted Dutch and Chinese children. The main research question is: (1) do the adopted Chinese children identify Chinese vocabulary better than the non-adopted Dutch children? A secondary research question is: (2) do the Chinese control children identify their native language vocabulary at a high standard of accuracy, and better than the adopted Chinese children and the Dutch controls? This secondary question evaluates the validity and reliability of the vocabulary test for the children of the age group tested in the present project. To test vocabulary knowledge, a picture-word matching task was used, and as in the perception and production studies, two experiments were conducted: one testing the Cantonese adoptees against the Cantonese controls and one group of Dutch controls, and the other testing the Mandarin adoptees against the Mandarin controls and another group of Dutch controls.

5.1 Experiment 1: Cantonese vocabulary

5.1.1 Method

5.1.1.1 Participants

Participants were the same 22 Cantonese adoptees, 23 Dutch controls, and 22 Cantonese controls, as described in Chapter 2 (section 2.1.1).

5.1.1.2 Materials

5.1.1.2.1 Speech materials

In each of the three vocabulary tests, 12 Cantonese words were tested, with a total number of 36 words (10 verbs and 26 nouns) (see Table A1 in Appendix A). The 36 Cantonese words were translations of 36 Mandarin words (see section 5.2.1.2) selected from the Mandarin CDI,

as described in Chapter 1. In addition, 18 Dutch words chosen from the Dutch CDI in (Zink & Lejaegere, 2002) were used for practice before each test (see Table A3 in Appendix A), with all 18 words for the pretest, and six words of the 18 for the intermediate test and the posttest, separately⁵⁹.

5.1.1.2.2 Picture materials

In each test, 24 pictures (12 for the targets and 12 for the distractors) were used, with a total number of 72 pictures. Additionally, 36 pictures were used for practice. For each trial, two pictures (one for the target and one for the distractor) were presented on the top of the computer screen horizontally, with a reasonable distance in between, taking up 3/5 of the computer space (Figure 22).

5.1.1.2.3 Video materials

Six motivation videos were used in total, with two for each test. In addition, another five motivation videos were used for practice, with three for the pretest, one for the intermediate test, and one for the posttest.

Similar to the motivation videos used in the perception study in Chapter 3, a single baby dinosaur or baby panda jumps on a seven-step stairway towards a gift box at the bottom corner of the screen (Figure 22). The motivation videos consisted of a series of six videos, with the baby animal beginning at one bottom corner of the screen and jumping one step closer, in each consecutive video, towards the gift box at the other bottom corner of the screen. The final motivation video of each series showed the baby animal arriving at the stair closest to the gift box, the box opening, and a unique gift popping out, accompanied by stars, a balloon and cheerful sound effects (Figure 22). Each motivation video was played for 1,000 ms. The motivation video was presented at the bottom of the screen horizontally, taking up 2/5 of the space.

⁵⁹ The same six Dutch words were used for practice in the intermediate test and the posttest.

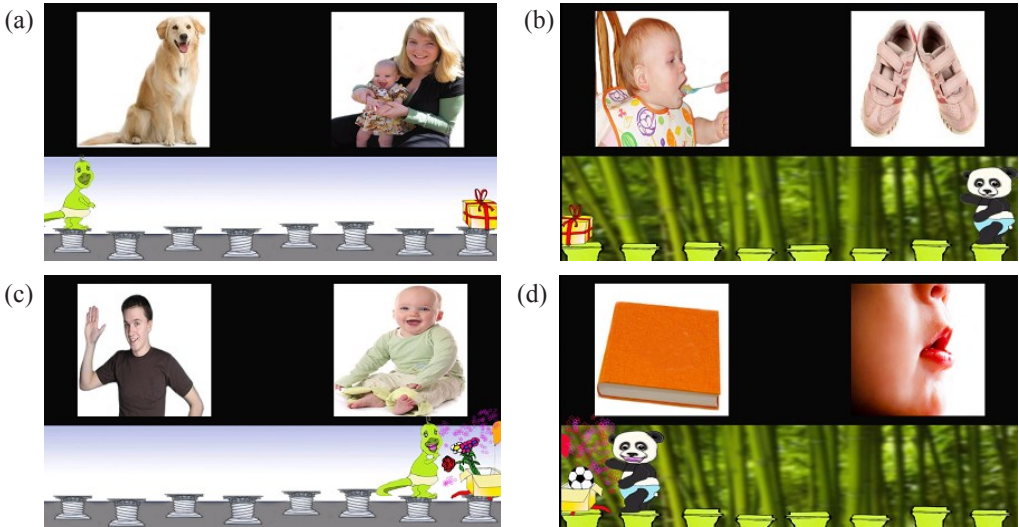


Figure 22 (a, b, c, d). Examples of video stills, each containing two pictures and a motivation video.

5.1.1.3 Design and procedure

Participants were tested with a picture-word matching task. They were asked to choose the picture which matched the word they heard from the computer⁶⁰, by pressing a corresponding button on a button box⁶¹. As in the perception tests in Chapter 3, there was no feedback about the correctness of the response, so the motivation videos, played after a response was given, served to motivate participants to continue the experiment. The task was again presented as a video game with sound, picture and video materials aligned.

Three vocabulary tests were conducted at the end of three experimental sessions, namely, the first, the second, and the fourth (i.e., last) session. Each test consisted of Dutch practice and Cantonese test trials. The pretest consisted of 18 practice and 12 test trials, the intermediate test of six practice and 12 test trials, and the posttest of six practice and 12 test trials. Each series of six trials formed a clear story line, by using six motivation videos showing the baby animal jumping on the stairs towards the gift box. Between the series of six trials, the initial position of the baby animal in the motivation videos changed from the left side to the right side of the screen. The pictures and motivation videos were presented

⁶⁰ Note that each word was played only once.

⁶¹ The left button on the button box represents the picture on the participants' left side of the screen, and the right button represents the picture on the participants' right side of the screen.

together on the screen (Figure 22), but with the motivation videos being played only after receiving a response from participants.

5.1.2 Results and discussion

Responses with RTs longer than 6,000 ms (123 responses, 5%) were considered as outliers, and discarded from analyses. Mauchly’s Test of Sphericity was used to assess whether the assumption of sphericity was violated, and if so, degrees of freedom were corrected using Greenhouse-Geisser estimates of Sphericity.

5.1.2.1 Adoptees vs. Dutch controls

Figure 23 and Table 17 show that there was no clear difference between the adoptees and the Dutch controls. Overall, the adoptees performed similarly to their Dutch control peers.

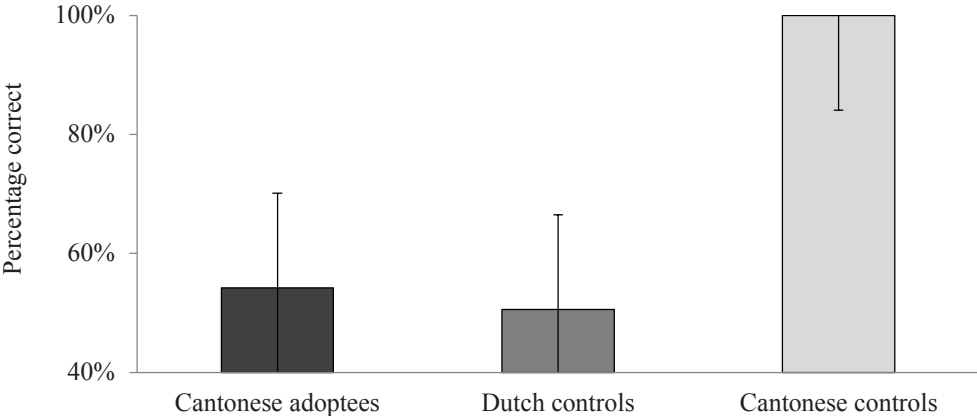


Figure 23. Percentage correct for the Cantonese adoptees, the Dutch controls, and the Cantonese controls averaged across three test instances. Error bars represent standard errors.

Table 17

Percentage correct for the Cantonese adoptees, the Dutch controls, and the Cantonese controls at each instance.

	Cantonese adoptees	Dutch controls	Cantonese controls
Pretest	44.9%	45.7%	100%
Intermediate test	55.2%	49.6%	100%
Posttest	62.7%	56.4%	100%

For the comparison between the adoptees and the Dutch controls, the same F_1 and F_2 ANOVAs were applied as in Chapter 3, with the proportions of correct responses, using the following variables⁶²: Group (Cantonese adoptees and Dutch controls), Test instance (Pretest, Intermediate test, and Posttest), and Parent Highest Education.

Results confirmed Figure 23 that there was no significant Group difference, $F_1(1, 42) = 1.790, p > .05, \eta^2_p = .041, F_2(1, 33) = 4.520, p < .05, \eta^2_p = .120$. Further, there was no effect of or interaction with Test instance, indicating that neither the adoptees nor the controls showed significant changes in their performance across the three test instances. The follow-up F_1 ANOVA including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates largely confirmed the results described above⁶³.

5.1.2.2 Adoptees and Dutch controls vs. Cantonese controls

Figure 23 and Table 17 also show that the Cantonese controls performed at ceiling across all test instances. ANOVAs including the Cantonese controls confirmed that these controls outperformed both the Cantonese adoptees and the Dutch controls (Cantonese controls versus Cantonese adoptees: $F_1(1, 42) = 995.892, p = .000, \eta^2_p = .960, F_2(1, 33) = 240.628, p = .000, \eta^2_p = .879$; Cantonese controls versus Dutch controls: $F_1(1, 43) = 522.499, p = .000, \eta^2_p = .924, F_2(1, 33) = 248.536, p = .000, \eta^2_p = .883$). As discussed in the perception study in Chapter 3, the highly accurate performance of the Cantonese controls affirmed both the feasibility of the task and the quality of the materials in the present study.

⁶² Group was a between-subjects (F_1) and within-items (F_2) variable, Test instance a within-subjects (F_1) and within-items (F_2) variable, and Parent Highest Education a covariate, because it was the only control variable that made the groups of children significantly different from each other (see section 2.1.1).

⁶³ The only differences were a significant effect of Age, $F_1(1, 37) = 4.130, p < .05, \eta^2_p = .100$, and a significant interaction between Test instance and Parent Highest Education, $F_1(2, 74) = 3.352, p < .05, \eta^2_p = .083$.

5.1.2.3 Residual memory of birth language vocabulary and Time Variables

Following the same procedures used in the perception and the production studies (Chapters 3 and 4), three correlational analyses were undertaken to investigate the effects of the three 'time variables' (Age, AoA, and LoR) on the Cantonese adoptees' residual memory of their birth language vocabulary (see Appendix F3 for the statistical analyses). Four task performance measures were used: Accuracy Overall, at the Pretest, at the Intermediate test, and at the Posttest.

Both Kendall's Tau correlations comparing each time variable against each of the four performance measures, as well as partial correlations between each performance measure and Age and AoA, controlling for LoR, failed to show any significant correlations at all. However, when the adoptees were split into the two sub-groups in which AoA and LoR were no longer confounded (see Table F.1.1 in Appendix F1), some significant effects emerged, though only for the larger sub-group, of 14 adoptees. For this group, (a) Age correlated positively with task performance at the posttest, i.e., older adoptees at the time of testing identified more Cantonese words there; (b) AoA was negatively correlated with performance at the pretest; adoptees who were younger at the time of adoption identified more Cantonese words in the pretest; and (c) LoR was positively correlated with pretest performance and overall accuracy; the longer the adoptees stayed in the Netherlands, the better they performed in identifying their birth language vocabulary at the pretest and overall. This effect of LoR is puzzling. However, it might still reflect an effect of Age to a certain extent, as discussed in Chapters 3 and 4. In the smaller sub-group, of 8 adoptees, no correlations appeared.

5.1.2.4 Effects of Siblings and China Visits

Effects on the vocabulary measure of Siblings and China Visits were investigated in two analyses as in Chapters 3 and 4. An F_1 ANOVA including Siblings and China Visits as covariates largely confirmed the main results presented in the main analyses⁶⁴ and showed no effect of or interaction with either Siblings or China Visits. A similar F_1 ANOVA adding

⁶⁴ In addition, there was a significant effect of Test instance, $F_1(2, 82) = 5.067, p < .01, \eta^2 p = .110$, with significantly higher scores in the posttest, compared to the pretest, and the intermediate test.

Siblings and China Visits as independent variables and leaving out all covariates likewise confirmed the main results⁶⁵, and showed no effect of or interaction with either variable.

5.1.2.5 Correlations between Dutch proficiency and the residual memory of birth language vocabulary

Kendall's Tau correlations also revealed no correlation between the adopted Cantonese children's proficiency in Dutch (as measured by the average Dutch CDI percentile scores) and their residual memory of Cantonese vocabulary (as measured by their average scores in the three Cantonese vocabulary tests).

5.1.2.6 Summary

The results of the present experiment suggest no conscious knowledge of birth language vocabulary in the Cantonese adoptees several years after adoption. Note that the group difference was significant in the F_2 analysis, but not in the F_1 analysis, which may be due to the large individual differences among the adoptees. Further, some effects of Age, AoA, and LoR were found, but only in the sub-group of 14 adoptees. There was no effect of Siblings or China Visits, either for the adoptees or for the controls, and no apparent relationship between Dutch proficiency and the residual memory of birth language vocabulary in the adoptees.

⁶⁵ Again, there was a significant effect of Test instance, $F_1(2, 76) = 9.039, p < .001, \eta_p^2 = .192$, with significantly higher scores in the posttest, compared to the pretest, and the intermediate test.

5.2 Experiment 2: Mandarin vocabulary

5.2.1 Method

5.2.1.1 Participants

Participants were the 26 Mandarin adoptees, 26 Dutch controls, and 18 Mandarin controls, as described in Chapter 2 (section 2.2.1).

5.2.1.2 Materials, design and procedure

As in Experiment 1 above, 36 words (10 verbs and 26 nouns) were used in the three vocabulary tests (12 per test) (see Table A2 in Appendix A). All 36 Mandarin words⁶⁶ were selected from the word list of Mandarin CDI in Hao, Shu, Xing & Li (2008), as described in Chapter 1. Additionally, the same 18 Dutch words as in Experiment 1 were used for practice, with all 18 words for the pretest, and six words out of 18 for the intermediate test and the posttest⁶⁷, respectively.

The picture and video materials, and the task, design and procedure were as in Experiment 1.

5.2.2 Results and discussion

As in Experiment 1, responses with RTs above 6,000 ms (130 responses, 5.1 %) were excluded from the main analyses. Mauchly's Test of Sphericity showed that the assumption of sphericity was not violated.

5.2.2.1 Adoptees vs. Dutch controls

Figure 24 and Table 18 again show no distinct difference between the adoptees and the Dutch controls. Both groups of children performed similarly in identifying (or rather: failing to identify) Mandarin words, across all three test instances.

⁶⁶ As described in Chapter 1, all 36 Mandarin words were reported to be understandable by 50% of native Mandarin children at the age of 12 months.

⁶⁷ Again, like in Experiment 1, the same six Dutch words were used for practice in the intermediate test and the posttest.

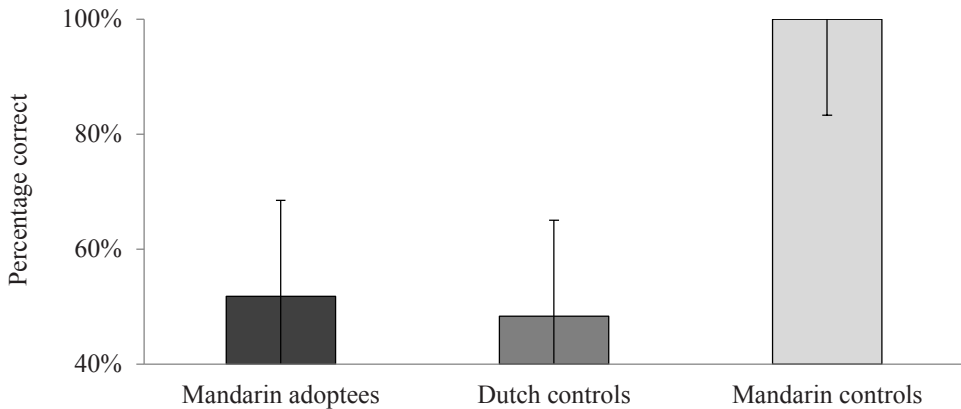


Figure 24. Percentage correct for the Mandarin adoptees, the Dutch controls, and the Mandarin controls across three test instances. Error bars represent standard errors.

Table 18

Percentage correct for the Mandarin adoptees, the Dutch controls, and the Mandarin controls at all three test instances.

	Mandarin adoptees	Dutch controls	Mandarin controls
Pretest	50.4%	48.3%	100%
Intermediate test	50.0%	47.5%	100%
Posttest	56.4%	49.8%	100%

ANOVAs as in Experiment 1 compared adoptees and controls with again proportions of correct responses as dependent variable, and Group (Mandarin adoptees and Dutch controls), and Test instance (Pretest, Intermediate test, and Posttest) as independent variables⁶⁸. No covariate was included since the adoptees and the controls were well matched on all control variables (see section 2.1.2).

Confirming Figure 24, there was no significant Group difference, $F_1(1, 50) = .810$, $p > .05$, $\eta^2_p = .016$, $F_2(1, 33) = 6.073$, $p < .05$, $\eta^2_p = .155$. Further, there was no effect of or interaction with Test instance, showing that no group of children changed their performance over time. The F_1 ANOVA including Age, Siblings, China Visits, Gender, Music, and Parent Highest Education as covariates confirmed this outcome.

⁶⁸ Group was a between-subjects (F_1) and within-items (F_2) variable, Test instance a within-subjects (F_1) and within-items (F_2) variable.

5.2.2.2 Adoptees and Dutch controls vs. Mandarin controls

Figure 24 and Table 18 also show that the Mandarin control children performed extremely well across three test instances. ANOVAs including the Mandarin controls showed that the Mandarin controls indeed outperformed both the Mandarin adoptees, $F_1(1, 42) = 303.874, p = .000, \eta^2_p = .879, F_2(1, 33) = 464.926, p = .000, \eta^2_p = .934$, and the Dutch controls, $F_1(1, 42) = 682.860, p = .000, \eta^2_p = .942, F_2(1, 33) = 412.191, p = .000, \eta^2_p = .926$.

5.2.2.3 Residual memory of birth language vocabulary and Time Variables

Kendall's Tau correlations between each time variable and each performance measure showed again no significant correlations, and nor did partial correlations between each performance measure and Age and AoA, controlling for LoR.

Finally, the division of the 26 adoptees into two sub-groups (to disentangle the correlation between AoA and LoR⁶⁹; see Table F.1.3 in Appendix F1) was exploited once more. Kendall's Tau correlations between each time variable and each performance measure in each sub-group showed a significant negative correlation between Age (and LoR) and task performance at the posttest, but only in the smaller sub-group of 6 Mandarin adoptees. The older these adoptees were at the time of testing (and the longer they had lived in the Netherlands), the less accurately they performed at the posttest. There were no other correlations.

5.2.2.4 Effects of Siblings and China Visits

As in Experiment 1, two analyses addressed the effects of Siblings and China Visits on the residual vocabulary memory, and again as in Experiment 1, these showed no effect of either of these variables.

5.2.2.5 Correlations between Dutch proficiency and the residual memory of birth language vocabulary

Kendall's Tau correlations between Dutch proficiency (average percentile scores on the Dutch CDI) and residual memory of vocabulary (average scores on the Mandarin vocabulary tests) showed, in contrast to Experiment 1, a significant negative correlation for the adopted

⁶⁹ Like in Chapter 3, Note that AoA and LoR were no longer correlated in either sub-group, but Age and LoR were still significantly correlated.

Mandarin children. Thus the higher the adoptees scored in the Dutch CDI, the lower they scored in the identification of their birth language vocabulary.

5.2.2.6 Summary

The results of Experiment 2 confirmed the conclusion of Experiment 1, namely that the Mandarin adoptees showed no conscious knowledge of their birth language vocabulary, across three test instances. Similar to the findings of Experiment 1, the group difference in Experiment 2 was also significant in the F_2 (items) analysis, but not in the F_1 (subjects) analysis, which might have resulted from individual differences among the adoptees. Contrary to Experiment 1 in which positive effects of Age and LoR were reported, in the present experiment negative effects of Age and LoR were found (only in the six adoptees who were adopted between 10 and 15 months old). Confirming Experiment 1, the present experiment showed no effect of either Siblings or China Visits. Finally, in contrast to Experiment 1, Experiment 2 showed a negative relationship between Dutch proficiency and the Mandarin adoptees' residual memory of their birth language vocabulary.

5.3 General discussion

5.3.1 No evidence of conscious knowledge of birth language vocabulary

Results in both Experiments 1 and 2 consistently showed no clear evidence of retained conscious knowledge of the birth language vocabulary in the Chinese adoptees, several years after adoption. This result is consistent with earlier studies which tested birth language vocabulary as an indication of knowledge of birth language in international adoptees, as discussed in Chapter 1 (e.g., Choi, 2014; Oh et al., 2010). Furthermore, the present result is largely in line with Isurin's (2000) report that a nine-year-old Russian girl gradually lost the production of her birth language vocabulary after adoption. However, the present result is in contrast to Schmid (2002) in which German Jews showed good maintenance of their L1 lexicon, even several decades after immigration. However, note that the German Jews may have probably received continuous exposure to German from their family members or relatives after immigration, contrary to the Chinese adoptees who had been completely cut off from their birth language after adoption. Together with previous studies, the present study confirmed that international adoptees forget their birth language after adoption, since the communicative function of the language is no longer required, as discussed in Chapter 1.

5.3.2 Effects of AoA and Age (and LoR)

The effects of AoA and Age (and LoR) were different in Experiments 1 and 2. For AoA, no effect was found in Experiment 2, though Experiment 1 showed a hint of evidence (in the sub-group of 14 adoptees) that adoptees who were adopted at an early age tended to identify their birth language vocabulary better at the pretest. This effect of AoA is counter-intuitive, because the adoptees who were adopted at an older age, not those who were adopted at an earlier age, were expected to have learned more birth language vocabulary, thus they were likely to have a better memory of their birth language vocabulary after adoption. However, note that the present effect of AoA was only based on a sub-group of 14 Cantonese adoptees, and was not confirmed in the Mandarin experiment; no generalization can really be made.

With respect to Age, both experiments showed positive effects at the posttest (though only in a sub-set of the data of each experiment). This result suggests that older adoptees were cognitively more mature, and thus capable of better performance after re-exposure at the posttest, as discussed in Chapter 4.

LoR produced different effects in the two experiments (again only for a sub-set). In Experiment 2 LoR positively affected posttest performance, while in Experiment 1 the effect of LoR on overall accuracy and on performance at the pretest was negative. The positive effect of LoR in Experiment 2 might be a reflection of the positive effect of Age, due to their inevitable correlation (see Chapter 3). The negative effect of LoR in Experiment 1 supports the common expectation that the longer the adoptees had been in the Netherlands, the less birth language vocabulary they retained. However, given the inconsistency, caution is called for when interpreting these results.

Within the Dutch control groups, different effects of Age were found in the two experiments: a positive effect of Age on overall accuracy in Experiment 1, but no Age effect in Experiment 2. It seems that older control children performed better overall. As discussed above, older children's greater cognitive maturity can produce better performance after training.

Variable effects of AoA, Age, and LoR were found in Experiments 1 and 2, but as discussed earlier, the number of adoptees in the present project is rather small for

correlational analyses. Future research needs to draw on a bigger sample size to study the effects of AoA, Age, LoR on the residual memory of birth language vocabulary.

5.3.3 Effects of Siblings and China Visits

Consistently, Experiments 1 and 2 showed no effect of either Siblings or China Visits. The participants with adopted siblings, and those who had visited China before test, did not perform differently from the participants without adopted siblings, and those who had never visited China. This result is different from that in the perception study (Chapter 3) where significant effects of these variables on the perception of Chinese sounds were found, particularly in the Mandarin adoptees and the Dutch controls. One possible account for the null effect in the present study is that the Chinese exposure from the adopted siblings and the visits in China might be too limited to elicit an effect on variables involving vocabulary.

5.3.4 Relationship between Dutch proficiency and the residual memory of birth language vocabulary

Experiments 1 and 2 showed different patterns of results for the relation between Dutch proficiency and the residual memory of birth language vocabulary: no correlation for the Cantonese adoptees, but a significant negative correlation for the Mandarin adoptees. This negative correlation suggests that international adoptees with well-maintained memory of their birth language may experience a delayed development of their adoptive language. Recall that Pollock, Chow, & Tamura (2004) found no evidence of cross-phonological interference between adoptive language and birth language. Interesting question arises for future research in that no study so far has systematically investigated the relationship between adoptive and birth languages.

This study has systematically investigated the residual memory of birth language vocabulary several years after adoption, in a large number of adopted Chinese children. The evidence consistently suggests no conscious knowledge of birth language vocabulary, although some evidence could indicate that older adoptees at the time of testing identified more birth language words after re-exposure at the posttest. Finally, the present data also provided some support for the suggestion that being less proficient in Dutch was related to better preserved birth language knowledge.

Chapter 6: General discussion

This dissertation is the first project that has systematically investigated the residual memory of birth language in a large number of internationally adopted children. Chinese children who were adopted from Cantonese and Mandarin Chinese speaking areas in China by Dutch speaking families were tested on the perception and the production of Chinese phonological contrasts through perceptual training. In addition, the adoptees' conscious knowledge of their birth language vocabulary was studied. The main findings confirm and expand earlier evidence that although international adoptees became less sensitive to their birth language sounds after adoption, and lose their conscious knowledge of their birth language vocabulary; this does not necessarily mean a total loss of the birth language memory. The adoptees were better at perceiving the phonological contrasts of their birth language after completing the perceptual training, compared to the control children who had had no prior experience with any Chinese language. Most interestingly, the adoptees produced the phonological contrasts of their birth language significantly better than the control children, before, between, and after perceptual training.

Below, I summarize and discuss the main findings of the dissertation in thematic sections, largely in accordance with those presented in the experimental chapters (Chapters 3, 4 and 5), and provide conclusions from the results.

6.1 Long-term benefit of early linguistic experience

The adopted Chinese children in the present project, conducted several years after their adoption, demonstrated a robust advantage over the Dutch control children in the perception (Chapter 3) and the production (Chapter 4) of Chinese phonological contrasts, although they showed no evidence of conscious knowledge of their birth language vocabulary (Chapter 5). These findings are in line with earlier work on adult heritage learners who overheard and/or spoke a language during their early childhood: early acquired phonological knowledge has long-term effects on relearning the language later in life (e.g., Production: Au et al., 2002, 2008; Knightly et al., 2003; Perception: Bowers et al., 2009; Oh et al., 2003; Tees & Werker, 1984). However, note that those adult heritage learners had been continuously exposed to their childhood language later in life (if only limited), as discussed in Chapter 1. In contrast, the adopted Chinese children in the present project had been completely cut off from their

birth language since adoption. Given that, the performance of the Chinese adoptees in the present project is striking.

With respect to the perception of birth language contrasts, both the Mandarin and the Cantonese adoptees outperformed the Dutch controls after finishing perceptual training, despite the fact that they did not differ from the controls before or between the perceptual training blocks. The present finding confirmed previous studies on adult and teenage adoptees that early acquired phonological knowledge of the birth language can be re-activated through re-exposure (Choi, 2014; Pierce et al., 2014; Singh et al., 2011), even though this knowledge has become difficult to retrieve before re-exposure (Pallier et al., 2003; Ventureyra et al., 2004). Taken together, the present finding provided valuable evidence for the existence of long-term retention of childhood language experience, particularly in the domain of speech perception.

With respect to the production of birth language contrasts, both the Mandarin and the Cantonese adoptees outperformed the Dutch controls before, between, and after perceptual training blocks. The present finding is in line with previous studies on heritage learners who had spoken a language briefly during their childhood and were able to produce the sounds of the language well after relearning, as discussed above. Importantly, the present finding is also consistent with the study on Korean adoptees by Choi (2014) in which adult Korean adoptees produced their birth language contrasts significantly better than the control participants after perceptual training. Most strikingly, the present result demonstrated such advantage in the production of Chinese contrasts for the adoptees even before perceptual training. As discussed in Chapter 4, the robust advantage of birth language production presented in the adopted Chinese children might be due to their shorter period of timespan of being isolated from their birth language, compared to those adult Korean adoptees in Choi (2014) who had been separated from their birth language for several decades. Thus, the present study expands the existing evidence of long-term benefit of childhood language experience in the domain of speech production.

Additionally, the adoptees' particular physiological and acoustic characteristics, and nearly-mature vocal tract structure may contribute to the high accuracy production of their birth language contrasts. Vocal tract configuration is crucial for human speech production (Vorperian et al., 2005). It was reported that adult Chinese speakers have significantly different oral volume and vocal tract volume, compared to White Americans and African

Americans (Xue, Hao, & Mayo, 2006). As for children, there is data showing that most of their vocal tract structures have reached between 55% and 80% of the adult size when they reach the age of 18 months (Vorperian et al., 2005). In the present study, the adoptees were adopted around an average age of two years, thus their vocal tract structures were expected to have been nearly mature at the time of adoption. This may have provided the adoptees with some advantage in producing the Chinese contrasts, even before the re-exposure through perceptual training.

6.2 Effects of AoA and Age (and LoR)

Research so far provided little hard evidence on the effects of AoA, Age, and LoR. In the present project, hints of effects of AoA, Age, and LoR were found across experiments, though highly variable. With respect to AoA, contrary to the null effect in Oh et al. (2010), Singh et al. (2011), and Choi (2014), the present project provides some suggestion of a positive effect of AoA in both the perception and the production of the Chinese contrasts, particularly in the Mandarin adoptees. The sub-group of 19 Mandarin adoptees who were adopted at a later age perceived their birth language contrasts better, particularly before perceptual training, and produced Mandarin tone 2 better. These results are similar to the finding of Hyltenstam et al. (2009) that a positive effect of AoA on the perception of birth language phonemes was observed in (only) two best-performing adoptees. The positive direction of the effect conforms to the common-sense expectation that the longer the adoptees had been exposed to their birth language, the better they may have preserved their birth language knowledge after adoption. Interestingly, this positive effect of AoA was not replicated in the Cantonese adoptees, which is likely to be caused by the small sample size and the bilingual situation in the Cantonese Chinese areas in China, as discussed in Chapter 3. However, in contrast to the non-effect in perception and production, a negative effect of AoA on the overall identification of the Cantonese vocabulary was reported for the sub-group of 14 Cantonese adoptees. The negative direction of the effect is counter-intuitive. Note that no consistent result was found in the complete groups of the Cantonese and the Mandarin adoptees, thus any interpretation of the present result should be made with caution.

With respect to Age and LoR, to the best of our knowledge no study so far has systematically investigated these two variables. In the present project, variable effects of Age and LoR in different directions were reported, across three experimental chapters. Overall, the effect of LoR largely mirrored the effect of Age, which is probably due to their linear

relation. Thus, the effects of Age and LoR were jointly referred to across chapters as the effects of Age (and LoR). *Positive effects:* Age (and LoR) was positively correlated with the adoptees' *perception* of their birth language contrasts, particularly after training. In addition, Age (and LoR) was positively correlated with the *production* of the affricate contrast for the Cantonese adoptees, and with the *production* of the tone contrast for the Mandarin adoptees, particularly after perceptual training. Furthermore, Age (and LoR) was positively correlated with the overall *identification of birth language vocabulary* for the Cantonese adoptees. Interestingly for the Dutch controls, Age was also reported to be positively correlated with several task performances, namely, *perception* of the Cantonese tone contrast, *production* of the Cantonese tone contrast at the posttest, *production* of the Mandarin unaspirated retroflex affricate, and the overall *identification of Cantonese vocabulary*. As discussed in the previous chapters, this positive effect of Age is likely to have resulted from older children's cognitive maturation; they were more task-ready, thereby able to achieve better performance through learning. *Negative effects:* Age (and LoR) was negatively correlated with the *production* of the affricate contrast for the Cantonese adoptees, and with the *identification of birth language vocabulary* at the posttest for the Mandarin adoptees. The reason for the negative effects is unclear. As discussed in the previous chapters, the sample size in the present project is rather small for correlational analyses. Further research is needed to systematically study the effects of the AoA and Age (and LoR) in a much bigger number of adoptees.

6.3 Effects of Siblings and China Visits

Similar to the effects of AoA and Age (and LoR), variable effects of Siblings and China Visits were found, particularly in the perception and the production studies. Note that null effect of either Siblings or China Visits was reported in the vocabulary study. With respect to Siblings, contrary to the null effect in Choi (2014), the present project showed several positive effects in both the perception and the production of the Chinese contrasts. In particular, the Dutch controls with adopted Cantonese siblings showed better *perception* of the Cantonese contrasts, and the children (both the adoptees and the controls) with adopted Mandarin siblings showed better *perception* of the Mandarin tone contrast. Further, the children (both the adoptees and the controls) with adopted Cantonese siblings showed better *production* of the Cantonese affricate contrast. As discussed in previous chapters, the positive effects of Siblings probably rest on two factors: re-exposure to the birth language sounds from the adopted siblings, and the high motivation influenced by the adopted siblings. As for

the scarce effect particularly for the Cantonese adoptees, it might be explained by the small number of Cantonese adoptees who had the opportunity to have overheard their birth language sounds from their adopted siblings, and the bilingual situation in the Cantonese speaking areas in China, as suggested in Chapter 3.

With respect to China Visits, the children (both the adoptees and the controls) who had visited the Mandarin speaking areas in China showed better *perception* of the Mandarin contrasts, and the children (both the adoptees and the controls) who had visited the Cantonese speaking areas in China showed better *production* of the Cantonese affricate contrasts particularly (only in the rating task). Further, the controls who had visited the Mandarin speaking areas in China showed similarly good *production* of the Mandarin tone contrast (only in the rating task), but those controls who had never visited China were outperformed by the adoptees who had also never visited China after adoption in the same task. All those positive effects of China Visits seem to suggest that exposure (re-exposure for the adoptees) to the Chinese sounds -though through a short visit in China- helps perceive and produce Chinese contrasts. As for the Cantonese adoptees and the Dutch controls in the perception study, no effect of China Visits was reported for the Cantonese adoptees, and a negative effect of China Visits for the controls. Such counter-intuitive results are probably caused by the dominant position of Mandarin Chinese in the Cantonese speaking areas in China, as discussed in Chapter 3.

6.4 Impacts of monolingual versus bilingual input

As discussed in previous chapters, Mandarin Chinese is the official language used in China, which led to the Cantonese speaking area becoming a bilingual region. Compared to the Mandarin adoptees who had been only exposed to Mandarin Chinese before adoption, the Cantonese adoptees were likely to have overheard both Cantonese Chinese and Mandarin Chinese before adoption. To certain extent, this difference in the birth language input may explain the different effects of AoA between the two groups of adoptees - whereas the Mandarin adoptees showed a positive effect of AoA on the perception of their birth language contrasts before perceptual training, no effect of AoA was found for the Cantonese adoptees due to the insufficient input of Cantonese before adoption. Furthermore, in contrast to the Cantonese adoptees who showed an improvement in the perception of their birth language contrasts only at the posttest (after finishing the perceptual training), the Mandarin adoptees

showed signs of improvement even at the intermediate test (albeit non-significantly). Such different patterns of improvement might also be caused by the different amount of birth language input. Finally, in contrast to the Mandarin adoptees who showed consistently positive effects of Siblings and China Visits, limited effects -only in the production of the Cantonese affricate contrast- were found for the Cantonese adoptees. It is likely that the children with adopted Cantonese siblings and the children who had visited the Cantonese speaking areas in China were exposed to a combination of Cantonese and Mandarin sounds, thus the exposure to Cantonese sounds was too restricted to have elicited a consistent effect. The interesting questions that arise concern how well the Cantonese adoptees may process Mandarin Chinese, and whether they may develop both birth languages equally well. As mentioned in Chapter 2, the present experiment was rather long for the young children tested in the present project; there was no space in the project to investigate both potential birth languages in the young Cantonese adoptees. Thus, the questions are open for future research to answer.

6.5 Conclusion

Even though non-continuous use may result in loss of a language (as predicted by the Activation Threshold Hypothesis: Paradis (2007)), the results in the present project suggests that it is not necessarily a complete loss. In this dissertation, two groups of Chinese adoptees demonstrated a robust advantage over those non-adopted Dutch controls in both the perception (Chapter 3) and the production (Chapter 4) of their birth language contrasts several years after adoption, although they did not show an evidence of conscious knowledge of the birth language vocabulary (Chapter 5). These findings confirm early evidence from teenage and adult adoptees that early acquired phonological knowledge of birth language has long-term effects despite non-continuous exposure, and that these effects assist both the perception and the production of sounds in the language.

Furthermore, the present results expand previous evidence with an early advantage of birth language production in the Chinese adoptees even before training. This underscores an important relationship between the timespan of being separated from the birth language and (particularly) the production of the birth language contrasts. In addition, the present study provides first hints of effects of bilingual versus monolingual birth language input, and effects of age of adoption, of age at testing, and of length of residence in the Netherlands on

birth language maintenance. These results imply that input may not be the sole important factor for birth language development; cognitive functions may have contributions as well. Further, re-exposure to birth language through adopted siblings or visits in China seems to have not only provided additional input of the language to the adoptees, but also increased their motivation to retrieve their birth language knowledge. Finally, there seems to be a cross-linguistic interference between the adoptive language and the birth language (particularly in vocabulary). All these effects contribute to filling gaps in the literature on birth language development in international adoptees. However, as discussed earlier, due to the small number of adoptees in the present project, the results reported above were inconsistent across analyses and experiments. Future research with a bigger sample size is required to further study what factors and how those factors contribute to birth language development in international adoptees.

References

- Au, T. K., Knightly, L. M., Jun, S.-A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science*, *13*, 238–243. doi:10.1111/1467-9280.00444
- Au, T. K., Oh, J. S., Knightly, L. M., Jun, S.-A., & Romo, L. F. (2008). Salvaging a childhood language. *Journal of Memory and Language*, *58*, 998–1011. doi:10.1016/j.jml.2007.11.001
- Bahrack, H. P. (1984). Fifty years of second language attrition: Implications for programmatic research. *The Modern Language Journal*, *68*, 105–118. doi:10.1111/j.1540-4781.1984.tb01551.x
- Bauer, R. S. (2010). LING6014 History and Structure of Cantonese Lecture 1: Introduction to Cantonese phonology. Retrieved from http://www.linguistics.hku.hk/cou/adv/ling6014/LING6014_Introduction_to_Cantonese_Phonology_0492010.pdf
- Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology* (Vol. 102). Walter de Gruyter.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, 13–34.
- Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. Retrieved from <http://www.fon.hum.uva.nl/praat/>
- Bouton, M. E. (1984). Differential control by context in the inflation and reinstatement paradigms. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 56–74.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*, 80–99.
- Bouton, M. E., & Peck, C. A. (1989). Context effects on conditioning, extinction, and reinstatement in an appetitive conditioning preparation. *Animal Learning & Behavior*, *17*, 188–198. doi:10.3758/BF03207634
- Bowers, J. S., Mattys, S. L., & Gage, S. H. (2009). Preserved implicit knowledge of a forgotten childhood language. *Psychological Science*, *20*, 1064–1069. doi:10.1111/j.1467-9280.2009.02407.x
- 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. doi:10.3758/BF03206911

- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English/r/and/l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*, 2299. doi:10.1121/1.418276
- Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *The Journal of the Acoustical Society of America*, *117*, 3890. doi:10.1121/1.1906060
- Chen, A., & Kager, R. (2012). Lexical Tone Perception in Early Infancy. In *Speech Prosody 2012*.
- Chen, C. J., Li, H., Shen, L., & Fu, G. (2001). Recognize tone languages using pitch information on the main vowel of each syllable. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on* (Vol. 1, pp. 61–64). IEEE. doi:10.1109/ICASSP.2001.940767
- Cheung, P. (1995). A speech therapy activity book on Hong Kong Cantonese phonology. *Hong Kong: Transaudio*.
- Ching, T. (1984). Lexical tone pattern learning in Cantonese children. *Language, Learning and Communication*, *3*, 243–414.
- Choi, J. (2014). *Rediscovering a forgotten language*. Unpublished doctoral dissertation, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands.
- Ciocca, V., & Lui, J. (2003). The development of the perception of Cantonese lexical tones. *Journal of Multilingual Communication Disorders*, *1*, 141–147.
- Cutler, A. (2012). *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge, Massachusetts London, England: The MIT Press.
- De Bot, K., Martens, V., & Stoessel, S. (2004). Finding residual lexical knowledge: The “Savings” approach to testing vocabulary. *International Journal of Bilingualism*, *8*, 373–382.
- De Bot, K., & Stoessel, S. (2000). In search of yesterday’s words: Reactivating a long-forgotten language. *Applied Linguistics*, *21*, 333–353.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 233–277.
- Gao, J., Shi, R., & Li, A. (2010). Categorization of lexical tones in Mandarin-learning infants. In *Speech Prosody* (pp. 946–949).
- Geren, J., Shafto, C. L., Snedeker, J., & Geren, J. (2009). Running fast but running behind: Language abilities in internationally-adopted children arriving between ages 2 and 9.
- Glennen, S. (2002). Language Development and Delay in Internationally Adopted Infants and Toddlers: A Review. *American Journal of Speech-Language Pathology*, *11*, 333.

- Glennen, S. (2005). New arrivals: Speech and language assessment for internationally adopted infants and toddlers within the first months home. In *Seminars in Speech and Language* (Vol. 26, pp. 10–21). Copyright© 2005 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA.
- Glennen, S., & Masters, M. G. (2002). Typical and atypical language development in infants and toddlers adopted from Eastern Europe. *American Journal of Speech-Language Pathology, 11*, 417–433. doi:10.1044/1058-0360(2012/11-0137)
- Gussenhoven, C. (1999). Tone systems in Dutch Limburgian dialects. In *symposium on cross-linguistic studies of tonal phenomena: Tonogenesis, typology, and related topics, Tokyo: Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Languages*.
- Gussenhoven, C., & Peters, J. (2008). De tonen van het Limburgs. *Nederlandse Taalkunde, 13*, 87–114.
- Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods, 40*, 728–733.
- Hardt, O., Wang, S.-H., & Nader, K. (2009). Storage or retrieval deficit: the yin and yang of amnesia. *Learning & Memory, 16*, 224–230. doi:10.1101/lm.1267409
- Hofer, S. B., Mrsic-Flogel, T. D., Bonhoeffer, T., & Hübener, M. (2006). Prior experience enhances plasticity in adult visual cortex. *Nature Neuroscience, 9*, 127–132. doi:10.1038/nm1610
- Hou, J. (2002). *现代汉语方言概论 (Introduction to Modern Chinese Dialects)*. 上海教育出版社 (Shanghai Educational Publishing House).
- Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H.-S. (2009). Dominant-language replacement: The case of international adoptees. *Bilingualism: Language and Cognition, 12*, 121–140. doi:10.1017/S1366728908004008
- Isurin, L. (2000). Deserted island or a child's first language forgetting. *Bilingualism: Language and Cognition, 3*, 151–166.
- 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. doi:10.1016/0010-0285(89)90003-0
- Knightly, L. M., Jun, S.-A., Oh, J. S., & Au, T. K. (2003). Production benefits of childhood overheard. *The Journal of the Acoustical Society of America, 114*, 465–474. doi:10.1121/1.1577560
- Knudsen, E. I. (1998). Capacity for plasticity in the adult owl auditory system expanded by juvenile experience. *Science, 279*, 1531–1533. doi:10.1126/science.279.5356.1531
- Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience, 16*, 1412–1425. doi:10.1162/0898929042304796

- Krakow, R., & Roberts, J. (2003). Acquisition of English vocabulary by young Chinese adoptees. *Journal of Multilingual Communication Disorders, 1*, 169–176.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience, 5*, 831–843. doi:10.1038/nrn1533
- Kwan, T.-W., Tang, W. S., Chiu, T.-M., Wong, L. Y., Wong, D., & Zhong, L. (2006). HUMANUM Chinese character database: A Chinese Character Database: With word-formations phonologically disambiguated according to the Cantonese dialect. The Chinese University of Hong Kong. Retrieved from <http://humanum.arts.cuhk.edu.hk/Lexis/lexi-can/>
- Law, Z. W. Y., & So, L. K. H. (2006). Phonological abilities of hearing-impaired Cantonese-speaking children with cochlear implants or hearing aids. *Journal of Speech, Language and Hearing Research, 49*, 1342–1353. doi:10.1044/1092-4388(2002/093)
- Lee, T., Kochanski, G., Shih, C., & Li, Y. (2002). Modeling tones in continuous Cantonese speech. In *INTERSPEECH*.
- Leung, M.-T., Law, S., & Fung, S.-Y. (2004). Type and token frequencies of phonological units in Hong Kong Cantonese. *Behavior Research Methods, Instruments, & Computers, 36*, 500–505. doi:10.3758/BF03195596
- Liu, C. Y. (2007). 18-25 个月儿童普通话的语音发展 (上海地区)(Mandarin phonetic development in 18-25 months old children: An investigation in Shanghai area). Unpublished doctoral dissertation, 上海师范大学(Shanghai Normal University).
- Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*, 251–285. doi:10.1017/S0272263100009165
- MacKain, K. S., & Best, C. T. (1981). Categorical perception of English /r/ and/l/by Japanese bilinguals. *Applied Psycholinguistics, 2*, 369–390. doi:10.1017/S0142716400009796
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition, 106*, 1367–1381.
- Meng, H., Zee, E., & Lee, W.-S. (2007). A contrastive phonetic study between Cantonese and English to predict salient mispronunciations by Cantonese learners of English. *Unpublished Article. The University of Hong Kong*.
- Nelson, T. O. (1978). Detecting small amounts of information in memory: Savings for nonrecognized items. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 453–468. doi:10.1037/0278-7393.4.5.453
- Nicoladis, E., & Grabis, H. (2002). Learning English and losing Chinese: A case study of a child adopted from China. *International Journal of Bilingualism, 6*, 441–454. doi:10.1177/13670069020060040401
- Norman, J. (2003). The Chinese dialects: phonology. In *The Sino-Tibetan Languages* (Vol. 3, pp. 72–83).

- Oh, J. S., Au, T. K.-F., & Jun, S. (2010). Early childhood language memory in the speech perception of international adoptees. *Journal of Child Language*, *37*, 1123. doi:10.1017/S0305000909990286
- Oh, J. S., Jun, S.-A., Knightly, L. M., & Au, T. K. (2003). Holding on to childhood language memory. *Cognition*, *86*, B53–B64. doi:10.1016/S0010-0277(02)00175-0
- Pallier, C., Dehaene, S., Poline, J.-B., LeBihan, D., Argenti, A.-M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex*, *13*, 155–161. doi:10.1093/cercor/13.2.155
- Paradis, J. (2007). Second language acquisition in childhood. In *Blackwell handbook of language development* (pp. 387–405). doi:10.1002/9780470757833.ch19
- Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of “interference” in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics*, *9*, 133–145.
- Pierce, L. J., Klein, D., Chen, J.-K., Delcenserie, A., & Genesee, F. (2014). Mapping the unconscious maintenance of a lost first language. *Proceedings of the National Academy of Sciences*, *111*, 17314–17319. doi:10.1073/pnas.1409411111
- Pollock, K., Chow, E., & Tamura, M. (2004). Phonological outcomes for preschoolers adopted from China as infants/toddlers. In *annual Child Phonology Conference, Tempe, AZ*.
- Pollock, K., Price, J., & Fulmer, K. (2003). Speech-language acquisition in children adopted from China: A longitudinal investigation of two children. *Journal of Multilingual Communication Disorders*, *1*, 184–193.
- Roberts, J. A., Pollock, K. E., Krakow, R., Price, J., Fulmer, K. C., & Wang, P. P. (2005). Language development in preschool-age children adopted from China. *Journal of Speech, Language and Hearing Research*, *48*, 93–107. doi:10.1044/1092-4388(2005/008)
- Rocha da Silva, P. (2006). La politique de l'enfant unique en République Populaire de Chine. (*The Politics of One Child in the People's Republic of China*) Université de Genève (*University of Geneva*), 22–28.
- Schmid, M. S. (2002). *First Language Attrition, Use and Maintenance: The case of German Jews in anglophone countries* (Vol. 24). John Benjamins Publishing Company.
- 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*, 949–959. doi:10.1111/j.1467-7687.2011.01044.x
- Snedeker, J., Geren, J. C., & Martin, I. (2005). The acquisition of English by internationally adopted preschoolers: a natural experiment in language development. Retrieved from http://www.wjh.harvard.edu/~lds/pdfs/Snedeker_Geren_Martin_2005.pdf

- Snedeker, J., Geren, J., & Shafto, C. L. (2007). Starting Over International Adoption as a Natural Experiment in Language Development. *Psychological Science, 18*, 79–87. doi:10.1111/j.1467-9280.2007.01852.x
- So, L. K. H., & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language, 22*, 473–496. doi:10.1017/S0305000900009922
- 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/Revue Canadienne de Psychologie, 38*, 579. doi:10.1037/h0080868
- Tsao, F. M. (2008). The effect of acoustical similarity on lexical-tone perception of one-year-old Mandarin-learning infants. *中華心理學刊 (Chinese Journal of Psychology), 50*, 111–124.
- Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *The Journal of the Acoustical Society of America, 120*, 2285.
- Tse, C.-Y. (1991). *The acquisition process of Cantonese phonology: A case study*. (Doctoral dissertation, The University of Hong Kong (Pokfulam, Hong Kong)).
- Tse, J. K.-P. (1978). Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language, 5*, 191–204. doi:10.1017/S0305000900007418
- Tse, S.-M. (1982). *The acquisition of Cantonese phonology*. University of British Columbia.
- Ventureyra, V. A. G., Pallier, C., & Yoo, H.-Y. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics, 17*, 79–91. doi:10.1016/S0911-6044(03)00053-8
- Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America, 117*, 338–350. doi:10.1121/1.1835958
- Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America, 113*, 1033–1043. doi:10.1121/1.1531176
- Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics, 37*, 35–44. doi:10.3758/BF03207136
- Wong, P., Schwartz, R. G., & Jenkins, J. J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language and Hearing Research, 48*, 1065–1079. doi:10.1044/1092-4388(2005/074)
- Wurm, S. A., K., B., T'sou, D. B., Li, R., Zhenghui, X., Zhenxing, Z., ... Dob (eds). (1987). *Language Atlas of China*. Hong Kong: Longman Group (Far East) Ltd. Pacific Linguistics Series C, (102).

- Xinhua Dictionary*. (2006). Beijing: The Commercial Press (in Chinese).
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61–83. doi:10.1006/jpho.1996.0034
- Xue, S. A., Hao, G. J. P., & Mayo, R. (2006). Volumetric measurements of vocal tracts for male speakers from different races. *Clinical Linguistics & Phonetics*, 20, 691–702. doi:10.1080/02699200500297716
- Zee, E. (1999). Chinese (Hong Kong Cantonese). In I. P. Association (Ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 58–60). Cambridge, UK: Cambridge University Press.
- Zhan, B., Cheung, Y.-S., Gan, Y.-E., (詹伯慧, 张日升, & 甘於恩). (1990). *珠江三角洲方言綜述 (A Survey of Dialects in the Pearl River Delta)* (Vol. 3). 新世紀出版社 (New Century Publishing House).
- Zhu, H. (2002). *Phonological development in specific contexts: Studies of Chinese-speaking children* (Vol. 3). Multilingual Matters.
- Zhu, H., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language*, 27, 3–42.
- Zink, I., & Lejaegere, M. (2002). N-CDI's Lijsten voor Communicatieve ontwikkeling: korte vormen. 2003; Aanpassing en hernormering van de Mac Arthur Short Form Vocabulary. *Acco: Leuven/Leusden*.

Appendix A

Chinese vocabulary selected from the MacArthur-Bates Communicative Development Inventory (tested in the picture-word matching task in Chapter 5)

Table A1

Cantonese Chinese vocabulary (in Guangzhou Cantonese Jyutping) at each test instance (with traditional Chinese characters and English translations)

Pretest	Intermediate test	Posttest
cong ⁴ 床 ⁶ (bed)	sik ⁶ fan ⁶ fan ⁶ 食飯 ⁶ (eat)	buk ⁹ nei ¹ nei ¹ 伏匿 ⁶ (peekaboo)
din ⁶ wa ⁶ 電話 ⁶ (telephone)	sek ³ sek ³ 錫 ³ (kiss)	a ² yi ⁴ 阿姨 ⁶ (aunt)
hêng ¹ ziu ¹ 香蕉 ⁶ (banana)	fa ¹ fa ¹ 花花 ⁶ (flower)	yem ² ye ⁵ 飲嘢 ⁶ (drink)
ngan ⁵ zei ² 眼仔 ⁶ (eye)	pen ⁴ guo ² 蘋果 ⁶ (apple)	mat ⁶ zei ² 襪仔 ⁶ (socks)
dan ² dan ² 蛋蛋 ⁶ (egg)	gao ³ gao ³ zü ¹ 覺覺 ⁶ 豬 ⁶ (sleep)	din ⁶ si ⁶ 電視 ⁶ (TV)
seo ² seo ² 手手 ⁶ (hand)	sü ¹ sü ¹ 書書 ⁶ (Book)	mun ⁴ 門 ⁶ (door)
gung ¹ zei ² 公仔 ⁶ (doll)	hai ⁴ zei ² 鞋仔 ⁶ (shoes)	sêu ² sêu ² 水水 ⁶ (water)
pou ⁵ pou ⁵ 抱抱 ⁶ (hug)	nai ⁵ nai ⁵ 奶奶 ⁶ (milk)	cê ¹ min ⁶ 出面 ⁶ (outside)
sam ¹ sam ¹ 衫衫 ⁶ (clothes)	wun ² wun ² 碗碗 ⁶ (bowl)	bêng ² bêng ² 餅餅 ⁶ (cookie)
tiu ³ mou ⁵ 跳舞 ⁶ (dance)	zêk ³ zei ² 雀仔 ⁶ (bird)	sei ² bak ⁶ bak ⁶ 洗白白 ⁶ (bathe)
siu ³ 笑 ⁶ (smile)	yi ⁵ zei ² 耳仔 ⁶ (ear)	co ⁵ 坐 ⁶ (sit)
fan ⁶ fan ⁶ 飯飯 ⁶ (food)	zêu ² zêu ² 嘴嘴 ⁶ (mouth)	béi ⁶ go ¹ 鼻哥 ⁶ (nose)

Note. The numbers represent different Cantonese tones, 1: Tone 1 (High-Level), 2: Tone 2 (High-Rising), 3: Tone 3 (Mid-Level), 4: Tone 4 (Low-Falling), 5: Tone 5 (Low-Rising), 6: Tone 6 (Low-Level), 7: Tone 7 (Entering High-Level), 8: Tone 8 (Entering Mid-Level), and 9: Tone 9 (Entering Low-Level).

Table A2

Mandarin Chinese vocabulary (in Pinyin) at each test instance (with simplified Chinese characters and English translations)

Pretest	Intermediate test	Posttest
chuang ² 床(bed)	chi ¹ fan ⁴ 吃饭(eat)	cang ² mao ¹ mao ¹ 藏猫猫(peekaboo)
dian ⁴ hua ⁴ 电话(telephone)	qin ¹ qin 亲亲(kiss)	a ¹ yi ² 阿姨(aunt)
xiang ¹ jiao ¹ 香蕉(banana)	hua ¹ 花(flower)	he ¹ 喝(drink)
yan ³ jing ¹ 眼睛(eye)	ping ² guo ³ 苹果(appel)	wa ⁴ zi 袜子(socks)
ji ¹ dan ⁴ 鸡蛋(egg)	shui ⁴ jiao ⁴ 睡觉(sleep)	dian ⁴ shi ⁴ 电视(TV)
shou ³ 手(hand)	shu ¹ 书(book)	men ² 门(door)
wa ² wa ² 娃娃(doll)	xie ² zi 鞋子(shoes)	shui ³ 水(water)
bao ⁴ bao ⁴ 抱抱(hug)	nai ³ 奶(milk)	wai ⁴ mian ⁴ 外面(outside)
yi ¹ fu ² 衣服(clothes)	wan ³ 碗(bowl)	bing ³ gan ¹ 饼干(cookie)
tiao ⁴ wu ³ 跳舞(dance)	xiao ³ niao ³ 小鸟(bird)	xi ³ zao ³ 洗澡(bathe)
xiao ⁴ xiao ⁴ 笑笑(smile)	er ³ duo 耳朵(ear)	zuo ⁴ 坐(sit)
fan ⁴ 饭(food)	zui ³ 嘴(mouth)	bi ² zi 鼻子(nose)

Note. The numbers represent different Mandarin tones, 1: Tone 1 (High-Level), 2: Tone 2 (High-Rising), 3: Tone 3 (Low-Dipping), and 4: Tone 4 (High-Falling).

Table A3

Dutch vocabulary for the practice (with English translations)

sap(juice)	fles(bottle)	slaap lekker(night night)
hallo(hello)	op(all gone)	auto(car)
hond(dog)	oma(grandma)	lekker(yum yum)
baby(baby)	papa(daddy)	poes(cat)
mama(mommy)	oh oh(uh oh)	bal(ball)
luier(diaper)	dag(bye)	beker(cup)

Note. All 18 words were used for practice in the pretest, and only the six words in the first column were used in the intermediate test and the posttest respectively.

Appendix B

Speech materials in Chapters 2 (Training), 3 (Perception) and 4 (Production).

Table B1

Cantonese speech materials (with international phonetic descriptions and traditional Chinese characters)

	Alveolar affricate contrast			Tone contrast		
	Unaspirated	Aspirated	Tone 2	Tone 5		
Perception test	a'zang [atsaŋ] "阿箏"	a'cang [ats ^h aŋ] "阿撑"	a'lün ² [alyŋ] "阿恋"	a'lün ⁵ [alyŋ] "阿嫵"		
	a'zi [atsi] "阿支"	a'ci [ats ^h i] "阿痴"	a'piu ² [apiu] "阿漂"	a'piu ⁵ [apiu] "阿票"		
	a'zên [atsɔŋ] "阿津"	a'cên [ats ^h ɔŋ] "阿春"	a'ying ² [ajɪŋ] "阿影"	a'ying ⁵ [ajɪŋ] "阿妖"		
	a'zai [atsai] "阿斋"	a'cai [ats ^h ai] "阿猜"	a'kuen ² [akw ^h en] "阿菌"	a'kuen ⁵ [akw ^h en] "阿君"		
	a'zou [atsou] "阿遭"	a'cou [ats ^h ou] "阿粗"	a'kuei ² [akw ^h ei] "阿鞋"	a'kuei ⁵ [akw ^h ei] "阿揸"		
	a'zei [atsɛi] "阿挤"	a'cei [ats ^h ɛi] "阿妻"	a'fen ² [atɛn] "阿粉"	a'fen ⁵ [atɛn] "阿愤"		
	a'zoi [atsɔi] "阿爻"	a'coi [ats ^h ɔi] "阿睬"	a'sü ² [asyu] "阿署"	a'sü ⁵ [asyu] "阿署"		
	a'za [atsa] "阿楂"	a'ca [ats ^h a] "阿叉"	a'wong ² [awɔŋ] "阿枉"	a'wong ⁵ [awɔŋ] "阿往"		
	a'zün [atsyn] "阿专"	a'cün [ats ^h yn] "阿穿"	a'wui ² [awui] "阿猥"	a'wui ⁵ [awui] "阿会"		
	a'zing [atsɪŋ] "阿睛"	a'cing [ats ^h ɪŋ] "阿蜻"	a'sé [asɛ] "阿舍"	a'sé [asɛ] "阿社"		
	a'zêu [atsoy] "阿狙"	a'cêu [ats ^h ɔy] "阿催"	a'pun ² [apun] "阿翻"	a'pun ⁵ [apun] "阿泮"		
	a'ziu [atsiu] "阿招"	a'ciu [ats ^h iu] "阿超"	a'yeo ² [ajeu] "阿黝"	a'yeo ⁵ [ajeu] "阿友"		
	Production practice	a'zao [atsau] "阿嘈"	a'cao [ats ^h au] "阿抄"	a'ma ² [ama] "阿嘛"	a'ma ⁵ [ama] "阿马"	
		a'zen [atsɛn] "阿珍"	a'cen [ats ^h ɛn] "阿衬"	a'yung ² [ajɔŋ] "阿拥"	a'yung ⁵ [ajɔŋ] "阿勇"	
Production test	a'zé [atɛ] "阿遮"	a'cé [ats ^h ɛ] "阿车"	a'tou ² [at ^h ou] "阿土"	a'tou ⁵ [at ^h ou] "阿肚"		
	a'zong [atsoŋ] "阿装"	a'cong [ats ^h ɔŋ] "阿仓"	a'séng ² [ascɛŋ] "阿想"	a'séng ⁵ [ascɛŋ] "阿上"		

Note. The numbers represent different Cantonese tones (see details in Appendix A1).

Table B2

Mandarin speech materials (with international phonetic descriptions and simplified Chinese characters)

	Retroflex affricate contrast		Tone contrast		
	Unaspirated	Aspirated	Tone 2	Tone 3	
Perception test	a 'zha [atʂa] "阿吒"	a 'cha [atʂʰa] "阿叉"	a 'luo² [aluo] "阿罗"	a 'luo³ [aluo] "阿裸"	
	a 'zhu [atʂu] "阿朱"	a 'chu [atʂʰu] "阿出"	a 'xi² [æi] "阿习"	a 'xi³ [æi] "阿喜"	
	a 'zhen [atʂən] "阿真"	a 'chen [atʂʰən] "阿琛"	a 'nian² [anien] "阿年"	a 'nian³ [anien] "阿撵"	
	a 'zhuang [atʂuɑŋ] "阿装"	a 'chuang [atʂʰuɑŋ] "阿窗"	a 'xuan² [æyen] "阿璇"	a 'xuan³ [æyen] "阿选"	
	a 'zhai [atʂai] "阿摘"	a 'chai [atʂʰai] "阿拆"	a 'pin² [apʰin] "阿贫"	a 'pin³ [apʰin] "阿品"	
	a 'zhou [atʂou] "阿周"	a 'chou [atʂʰou] "阿抽"	a 'liang² [alɪɑŋ] "阿良"	a 'liang³ [alɪɑŋ] "阿两"	
	a 'zhui [atʂuei] "阿追"	a 'chui [atʂʰuei] "阿吹"	a 'huan² [axuan] "阿环"	a 'huan³ [axuan] "阿缓"	
	a 'zhuo [atʂuo] "阿卓"	a 'chuo [atʂʰuo] "阿戳"	a 'miao² [amiau] "阿苗"	a 'miao³ [amiau] "阿秒"	
	a 'zhan [atʂan] "阿詹"	a 'chan [atʂʰan] "阿掺"	a 'ying² [ajɪŋ] "阿盈"	a 'ying³ [ajɪŋ] "阿影"	
	a 'zhun [atʂun] "阿淳"	a 'chun [atʂʰun] "阿春"	a 'ran² [aran] "阿然"	a 'ran³ [aran] "阿染"	
	a 'zhuai [atʂuai] "阿拽"	a 'chuai [atʂʰuai] "阿揣"	a 'niu² [aniou] "阿牛"	a 'niu³ [aniou] "阿纽"	
	a 'zheng [atʂəŋ] "阿争"	a 'cheng [atʂʰəŋ] "阿撑"	a 'shui² [aʂuei] "阿谁"	a 'shui³ [aʂuei] "阿水"	
	Production practice	a 'zhe [atʂə] "阿遮"	a 'che [atʂʰə] "阿车"	a 'yu² [ajy] "阿雨"	a 'yu³ [ajy] "阿雨"
	Production test	a 'zhong [atʂuŋ] "阿中"	a 'chong [atʂʰuŋ] "阿冲"	a 'hao² [axau] "阿豪"	a 'hao³ [axau] "阿好"
		a 'zhao [atʂau] "阿招"	a 'chen [atʂʰau] "阿超"	a 'mei² [amei] "阿梅"	a 'mei³ [amei] "阿美"
		a 'zhuan [atʂuan] "阿专"	a 'chuan [atʂʰuan] "阿川"	a 'tian² [atʰien] "阿甜"	a 'tian³ [atʰien] "阿舔"

Note. The numbers represent different Cantonese tones (see details in Appendix A2).

Table B3

Dutch speech materials used for perception practice (with international phonetic descriptions)

Contrast between /f/ and /t/	
afaan [afa:n]	ataan [ata:n]
afarg [afarχ]	atarg [atarχ]
afeg [afεχ]	ateg [atεχ]
afeits [afeits]	ateits [ateits]
afesk [afesk]	atesk [atesk]
afeuf [afø:f]	ateuf [atø:f]
afijp [afeip]	atijp [ateip]
afoes [afus]	atoes [atus]
afom [afɔm]	atom [atɔm]
afook [afɔ:k]	atook [ato:k]

Appendix C

Perceptual training: Analyses and Results

C.1 Cantonese

Following the same RT cut-off value as that of Experiment 1 in Chapter 3, responses with RTs longer than 10,000 ms (1076 responses (3.3% out of total)) were discarded from the main analyses. Mauchly's Test of Sphericity was used to assess whether the assumption of sphericity was violated, and if so, degrees of freedom were corrected using Greenhouse-Geisser estimates of Sphericity.

C.1.1 Adoptees vs. Dutch Controls

As Figure 8 (see Chapter 2) shows, there was no difference between the Cantonese adoptees and the Dutch controls, across all ten training blocks. The two groups of children performed similarly in their perception of the Cantonese contrasts. In addition, neither group of children showed significant changes in their performance over time.

For the comparison between the adoptees and the Dutch controls, the same ANOVAs as those in Chapter 3 were applied, with the proportions of correct responses as dependent variable, Group (Cantonese adoptees and Dutch controls), Training block (block 1 to 10), and Contrast type (Affricates and Tones) as independent variables, and Parent Highest Education as covariate (since that was the only control variable in which the two groups of children significantly differ from each other, see section 2.1.1).

Confirming Figure 8, results showed no effect of Group ($ps > .05$), or Training block ($ps > .05$). Additionally, there was no effect of or interaction with Contrast type, suggesting that the two groups of children perceived the Cantonese affricate and tone contrasts in a similar manner. Therefore, the two types of contrasts were collapsed in Figure 8; for a separate graph of each contrast, see Figure C.1a and C.1b below. Finally, there was no effect of or interaction with Parent Highest Education, $ps > .05$. An F_1 ANOVA similar to the one

described above, but including Age, Siblings, To China, Gender, Music, and Parent Highest Education as covariates, confirmed the main findings described above⁷⁰.

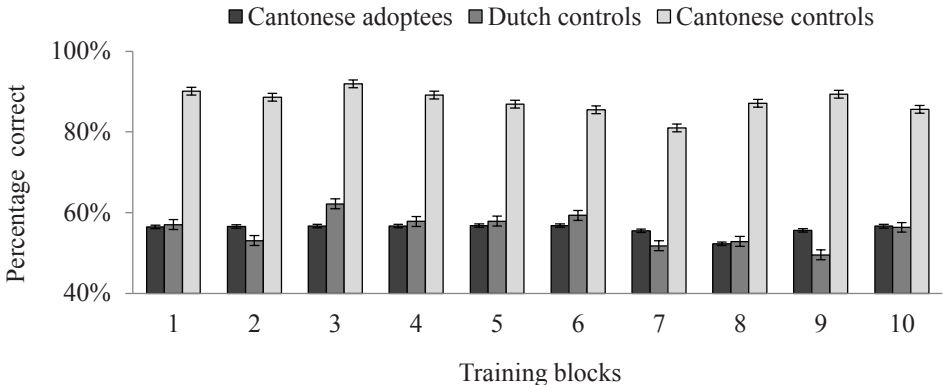


Figure C.1a. Percentage correct of the Cantonese affricate contrast for the Cantonese adoptees, the non-adopted Dutch controls, and the non-adopted Cantonese controls. Error bars represent standard errors.

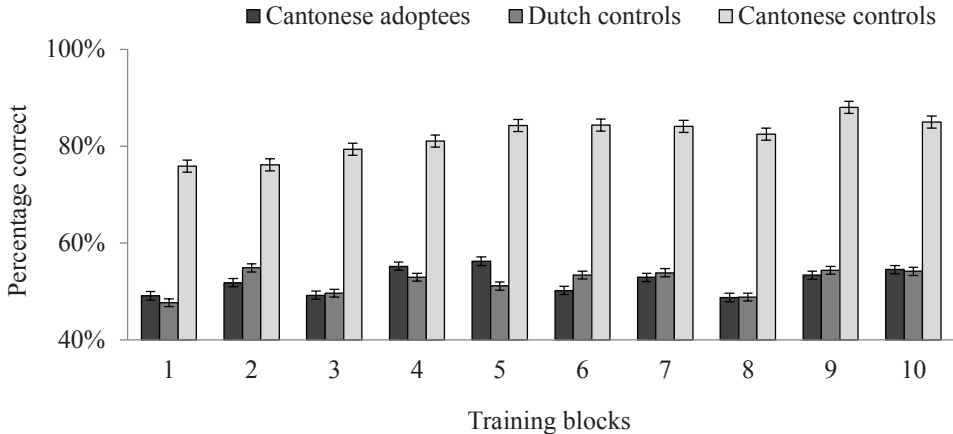


Figure C.1b. Percentage correct of the Cantonese tone contrast for the Cantonese adoptees, the non-adopted Dutch controls, and the non-adopted Cantonese controls. Error bars represent standard errors.

⁷⁰ Additionally, there was a significant three-way interaction among Training block, Contrast type, and Gender, $F_1(9, 333) = 2.569, p < .01, \eta^2_p = .065$.

C.1.2 Adoptees and Dutch Controls vs. Cantonese controls

As Figure 8 also shows, the Cantonese controls performed very well in all ten training blocks. Similar ANOVAs to those used in the main analyses, but including the Cantonese controls showed that the Cantonese controls outperformed both the Cantonese adoptees and the Dutch controls across all training blocks (Cantonese controls versus Cantonese adoptees: $F_1(1, 42) = 103.232, p < .001, \eta^2_p = .711$; $F_2(1, 2) = 214.547, p < .01, \eta^2_p = .991$; Cantonese controls versus Dutch controls: $F_1(1, 43) = 119.758, p < .001, \eta^2_p = .736$; $F_2(1, 2) = 169.535, p < .01, \eta^2_p = .988$).

C.2 Mandarin

Using the same RT cut-off value as that of Experiment 2 in Chapter 3, 689 responses (2.1% out of total) with RTs longer than 13,000 ms were excluded from the analyses. Mauchly's Test of Sphericity showed that the assumption of Sphericity was not violated.

C.2.1 Adoptees vs. Dutch Controls

Figure 9 (see Chapter 2) also shows no difference between that the Mandarin adoptees and the Dutch controls, across all training blocks. Further, none of the children in either group showed significant improvement in their performance over time.

Like in section 2.4.3.1.1, the same ANOVAs were conducted to compare the Mandarin adoptees and the Dutch controls. Again, the proportions of correct responses were used as dependent variable, Group (Mandarin adoptees and Dutch controls), Training block (block 1 to 10), and Contrast type (Affricates and Tones) as independent variables. Different from section 2.4.3.1.1, no covariate was included, since the Mandarin adoptees and the Dutch controls were well-matched in all control variables (see section 2.1.2).

Results confirmed Figure 9 that there was neither a significant Group difference⁷¹, $F_1(1, 50) = 4.612, p < .05, \eta^2_p = .084$, $F_2(1, 2) = .055, p > .05, \eta^2_p = .027$, nor a main effect of Training block ($ps > .05$). Additionally, there was no effect of or interaction with Contrast type ($ps > .05$), thus affricate and tone contrasts were collapsed in Figure 9; separate figures for the two contrasts were presented in Figure C.2a and C.2b below. A similar F_1 ANOVA to

⁷¹ Although the Group difference was significant across subjects (F_1), it was not significant across items (F_2), which was likely caused by the great variations in the training items.

that presented above, but including Age, Siblings, To China, Gender, Music, and Parent Highest Education as covariates, largely confirmed the main results described above⁷².

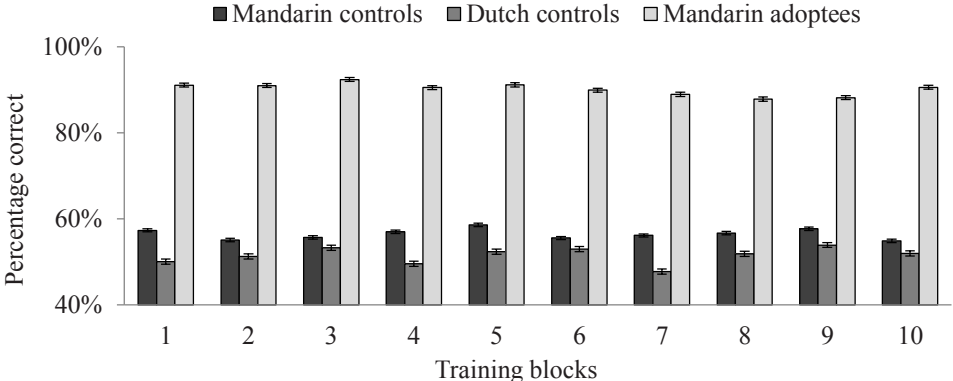


Figure C.2a. Percentage correct of the Mandarin affricate contrast for the Mandarin adoptees, the non-adopted Dutch controls, and the non-adopted Mandarin controls. Error bars represent standard errors.

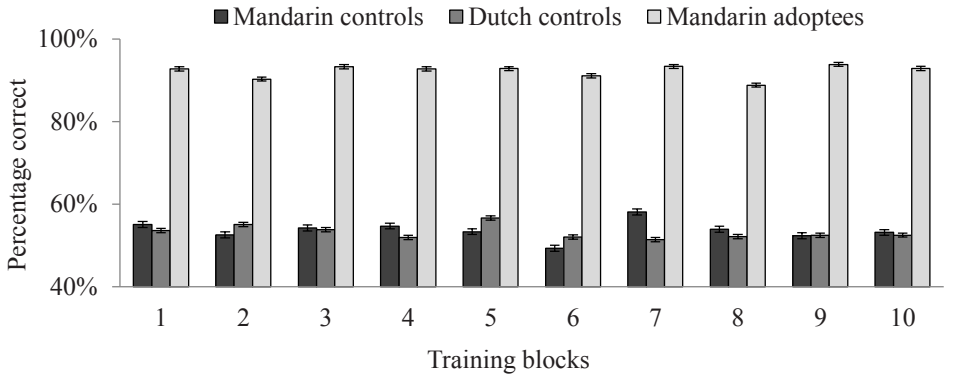


Figure C.2b. Percentage correct of the Mandarin Tone contrast for the Mandarin adoptees, the non-adopted Dutch controls, and the non-adopted Mandarin controls. Error bars represent standard errors.

⁷² In addition, significant interactions between Contrast type and Group ($F_1(1, 446) = 6.123, p < .05, \eta_p^2 = .122$), between Contrast type and Siblings, ($F_1(1, 44) = 4.273, p < .05, \eta_p^2 = .089$), between Training block and ToChina ($F_1(9, 396) = 1.926, p < .05, \eta_p^2 = .042$), and a significant effect of Age ($F_1(1, 44) = 9.034, p < .01, \eta_p^2 = .170$) were found. Follow-up analysis on the interaction between Contrast type and Group showed a significant Group difference for the affricate contrast: The adoptees perceived the affricate contrast better than the controls. Follow-up analyses on the interaction between Contrast type and Siblings showed a significant effect of Siblings for the affricate contrast: The children without adopted siblings performed better than those with adopted siblings. Follow-up analyses on the interaction between Training block and ToChina showed a significant effect of ToChina in Training blocks 1, 5, and 8: The children who had visited the Mandarin areas in China perceived the Mandarin contrasts better than those who did not, particularly in the training blocks 1, 5 and 8.

C.2.2 Adoptees and Dutch Controls vs. Mandarin controls

Figure 9 also shows a high-accuracy performance of the Mandarin controls across all training blocks. Similar ANOVAs to those used in the main analysis (section 2.4.3.2.1) including the Mandarin controls confirmed that the Mandarin controls performed significantly better than both the Mandarin adoptees and the Dutch controls across all training blocks (Mandarin controls versus Mandarin adoptees: $F_1(1, 42) = 282.494, p < .001, \eta^2_p = .871$; $F_2(1, 2) = 236.866, p < .01, \eta^2_p = .992$; Mandarin controls versus Dutch controls: $F_1(1, 42) = 503.754, p < .001, \eta^2_p = .923$; $F_2(1, 2) = 893.956, p < .01, \eta^2_p = .998$).

Appendix D

Questionnaire on participants' characteristics (in Dutch)

D.1 Questionnaire for Dutch control children

Vragenlijst-deel 1

I. Algemeen (gegevens kind)

Roepnaam: _____

Achternaam: _____

Jongen/Meisje: _____

Geboortedatum: _____ - _____ - _____

Dus de leeftijd van uw kind is nu +/-: _____

Wordt uw kind thuis alleen Nederlandstalig opgevoed of ook met een andere moedertaal?

Alleen Nederlands / Meertalig

Indien meertalig, welke talen? _____

Spreekt uw kind nog andere talen (ook al is het maar een beetje)? **JA / NEE**

Zo ja, welke talen? _____

II. School

Naar wat voor school gaat uw kind?

- Basisonderwijs: _____
- Anders, nl: _____

In welke groep zit uw kind? _____

Heeft uw kind ooit een jaar overgedaan? **JA /NEE**

Zo ja, welke groep? _____

III. Muziek vaardigheden

Krijgt/Kreeg uw kind muzieklessen (zang of instrument) naast de eventuele lessen op school?

JA / NEE

Zo ja, vul de onderstaande kolom in.

Muziekles / ervaring	Sinds	Hoe lang?

Vragenlijst-deel 2

I. Nederlandse familie

Wat is het hoogst genoten opleidingsniveau van u en uw partner?

- Vader: **Middelbare school / MBO / HBO / WO**
- Moeder: **Middelbare school / MBO / HBO / WO**

Heeft u nog andere kinderen die geadopteerd zijn uit China? **JA / NEE**

Zo ja, vul de informatie over uw andere adoptiekind in, in de onderstaande tabel.

Naam van het kind	Huidige leeftijd	Leeftijd op moment van adoptie	Naam van het kindertehuis	Regio van het kindertehuis in China (stad en/of provincie)

D.2 Questionnaire for Chinese adoptees

Vragenlijst-deel 1

I. Algemeen (gegevens kind)

Roepnaam: _____

Achternaam: _____

Jongen/Meisje: _____

Geboortedatum: _____ - _____ - _____

Dus de leeftijd van uw kind is nu +/-: _____

Wordt uw kind thuis alleen Nederlandstalig opgevoed of ook met een andere moedertaal?

Alleen Nederlands / Meertalig

Indien meertalig, welke talen? _____

Spreekt uw kind nog andere talen (ook al is het maar een beetje)? **JA** / **NEE**

Zo ja, welke talen? _____

II. School

Naar wat voor school gaat uw kind?

- Basisonderwijs: _____
- Anders, nl: _____

In welke groep zit uw kind? _____

In welke groep is uw kind na de adoptie begonnen? _____

Heeft uw kind ooit een jaar overgedaan? **JA** / **NEE**

Zo ja, welke groep? _____

III. Muziek vaardigheden

Krijgt/Kreeg uw kind muzieklessen (zang of instrument) naast de eventuele lessen op school?

JA / NEE

Zo ja, vul de onderstaande kolom in.

Muziekles / ervaring	Sinds	Hoe lang?

Vragenlijst-deel 2

I. Nederlandse familie

Wat is het hoogst genoten opleidingsniveau van u en uw partner?

- Vader: **Middelbare school** / **MBO** / **HBO** / **WO**
- Moeder: **Middelbare school** / **MBO** / **HBO** / **WO**

Heeft u nog andere kinderen die ook geadopteerd zijn uit China? **JA / NEE**

Zoja, zijn uw kinderen tegelijk geadopteerd uit en uit hetzelfde kindertehuis?

Vul de informatie over uw andere adoptiekind in, in de onderstaande tabel.

Naam van het kind	Huidige leeftijd	Leeftijd op moment van adoptie	Naam van het kindertehuis	Regio van het kindertehuis in China (stad en/of provincie)

Vragenlijst-deel 3

I. Adoptie-geschiedenis van uw kind

Hoe oud was uw kind toen hij/zij naar Nederland kwam? _____

Weet u in welk kindertehuis uw kind woonde voordat hij/zij naar Nederland kwam (in welke stad)? _____

Weet u waar uw kind was voordat hij/zij in het kindertehuis kwam? (Bv. bij biologische familie, in andere opvang, etc?) _____

- Woonden uw kind en zijn/haar biologische familie in hetzelfde dorp /dezelfde stad waar ook het kindertehuis was? **JA / NEE**

Zo ja, voor hoe lang woonde uw kind daar? _____

Zo nee, in welk(e) dorp/stad woonde uw kind samen met zijn/haar biologische familie? _____

voor hoe lang woonde uw kind daar? _____

- Als het kindertehuis niet in hetzelfde/dezelfde dorp/stad was, wat was de naam van het kindertehuis waarin uw kind woonde? _____

In welk(e) dorp/stad/provincie in China was dat kindertehuis? _____

Wanneer ging uw kind naar het kindertehuis? _____

Wanneer verliet uw kind het kindertehuis? _____.

- Als uw kind in verschillende kindertehuizen heeft gezeten, vult u dan de onderstaande tabel in.

Naam kindertehuis	Locatie kindertehuis (stad en/of provincie)	Vanaf wanneer	Tot wanneer

II. Oorspronkelijk taal (Chinees)

Kunt u de taalvaardigheden van de oorspronkelijke (Chinese) taal van uw kind op verschillende momenten aangeven? U kunt dit doen door kruisjes te zetten in de onderstaande tabel.

Kun u aangeven hoe vaak uw kind sprak/spreekt in zijn/haar oorspronkelijk taal op verschillende momenten in zijn/haar leven? U kunt dit met een kruisje aangeven in de onderstaande tabel.

Periode \ Frequentie	Spreken						
	Elke dag	Een paar keer per week	Eens per week	Een paar keer per maand	Eens per maand	Minder dan eens per maand	Nooit
Op moment van adoptie							
Tegenwoordig							

Spreekt u, uw partner of iemand in uw familie Chinees? **JA / NEE**

*Zo ja, welke Chinese taal? **Mandarijn / Kantonees***

Spreeken u en uw partner Chinees tegen uw kind? **JA / NEE**

Luistert of kijkt uw kind wel eens naar Chinese liedjes of (teken)films? **JA / NEE**

Zo ja, vul de onderstaande tabel in.

Soorten \ Frequentie	Hoe vaak						
	Elke dag	Een paar keer per week	Eens per week	Een paar keer per maand	Eens per maand	Minder dan eens per maand	Nooit
Chinese muziek							
Chinese (teken) films							

Heeft uw kind ooit Chinees les gehad, of zelf Chinees geprobeerd te leren? **JA / NEE**

Zo ja, graag toelichten. _____

Komt uw kind op een andere manier wel eens in contact met de Chinees taal? Bv. Heeft uw kind de mogelijkheden om mensen te ontmoeten die Chinees spreken? **JA / NEE**

Zo ja, vul de onderstaande tabel in.

Wie en waar? \ Frequentie	Hoe vaak					
	Elke dag	Een paar keer per week	Eens per week	Een paar keer per maand	Eens per maand	Minder dan eens per maand

Is uw kind sinds zijn/haar adoptie wel eens terug geweest in China?

	Jaar	Plaatsen bezocht	Duur van het verblijf	Doel, bv. roots reis, vakantie, werk	Heeft uw kind daar Chinese les gehad ? (Ja /Nee)
Eerste bezoek					
Tweede bezoek					
Derde bezoek					

Appendix E

Separate figures for the percentage correct of each phonological contrast in Chapter 3. Error bars represent standard errors.

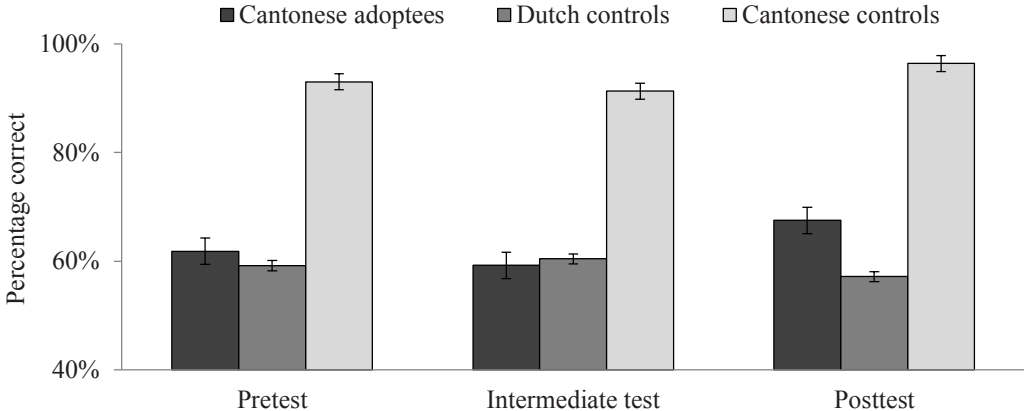


Figure E.1. Percentage correct of the Cantonese alveolar affricate contrast.

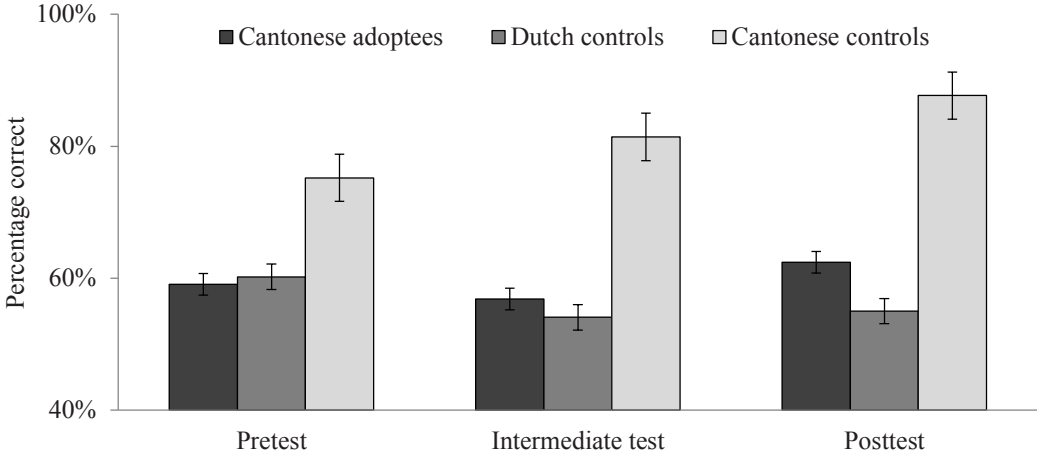


Figure E.2. Percentage correct of the Cantonese tone contrast.

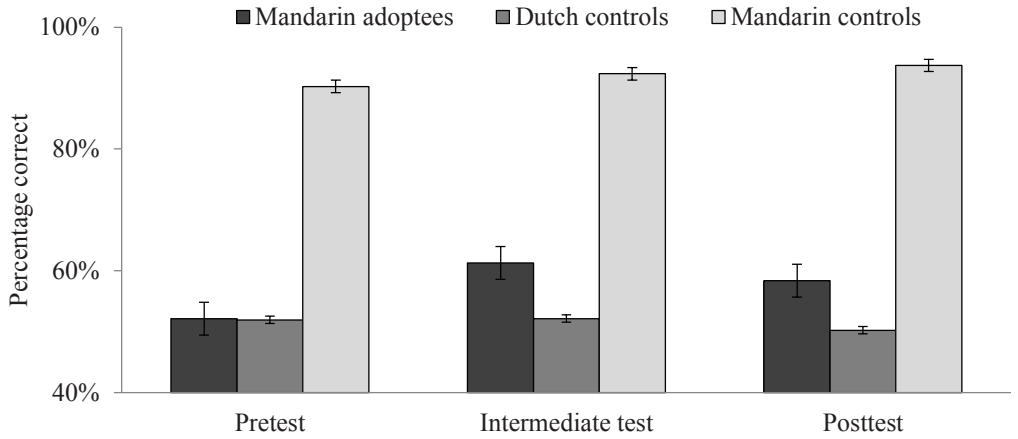


Figure E.3. Percentage correct of the Mandarin retroflex affricate contrast.

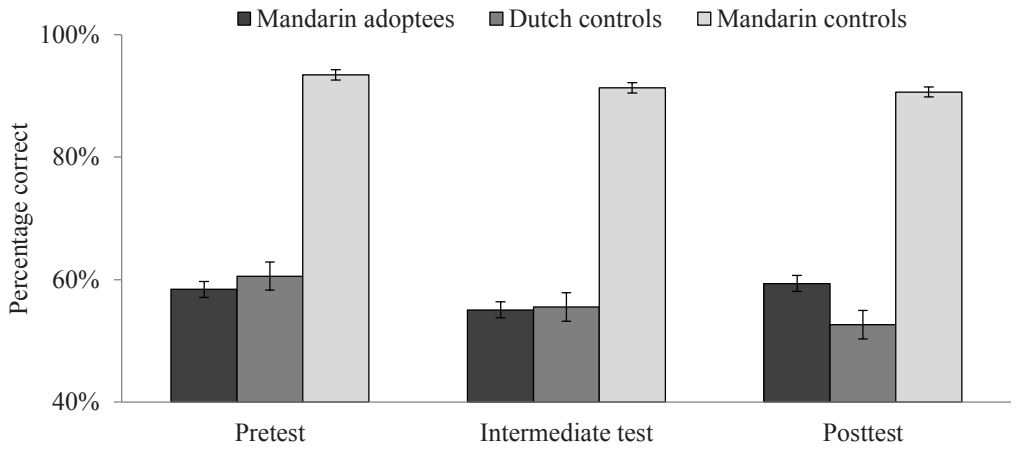


Figure E.4. Percentage correct of the Mandarin tone contrast.

Appendix F

Statistical analyses for the relationship among birth language development and Age, AoA, and LoR

F.1 Relationship among perception and Age, AoA, and LoR

This section details the analyses (Kendall's Tau correlation) for the effects of the three 'time variables' (Age, AoA, and LoR) on the perception of the Chinese contrasts. As described in Chapter 3, Age, AoA, and LoR, are linearly related ($Age = AoA + LoR$), and the correlation between Age and LoR is non-disentangled due to the fact that the longer the Cantonese adoptees stayed in the Netherlands, the older they became. As for the task performance, the same six measures as those described in Chapter 3 were used, namely, Accuracy Overall, at the Pretest, at the Intermediate test, at the Posttest, for Affricates, and for Tones. Three correlational analyses were applied for each language experiment.

F.1.1 Cantonese

Within the group of Cantonese adoptees, LoR was reported to be significantly correlated to both Age ($r = .880, p < .01$), and AoA ($r = -.657, p < .01$). Firstly, Kendall's Tau correlation between each time variable and each task performance measure for the 22 adoptees were carried out. As Table F.1.1a shows, no significant correlation was found.

Secondly, partial correlations among each measure of task performance and Age and AoA for the 22 adoptees, while controlling LoR confirmed previous analysis that there was no significant correlation (see Table F.1.1b). Note that the results of Age and AoA were identical, due to the linear relationship among the three time variables.

Finally, due the correlation between AoA and LoR can be disentangled by regrouping the 22 adoptees, we divided the 22 adoptees into two different sub-groups (see Table F.1.1c) based on their age, in such a way that LoR and AoA were not significantly correlated within each of the sub-groups. The first sub-group contained 14 children who were adopted between 19 and 54 months (i.e., 1;7-4;6 years); the second sub-group contained 8 children who were adopted at the age of 9-16 months (i.e., 0;9-1;4 years). Kendall's Tau correlational analyses between each of the time variables and each of the performance measures were conducted for each sub-group of adoptees. As Table F.1.1d shows, only in the sub-group of 14 adoptees,

Age showed significant positive correlations with several performance measures, i.e., overall perception of the Cantonese contrasts, perception of the contrasts at the Posttest, and perception for both the affricate and the tone contrasts. Similarly, LoR showed partially overlapping positive correlations as Age did, namely, with overall perception of the Cantonese contrasts, and perception of the contrasts at the Posttest. However, within the same sub-group, no correlation with AoA was found. Finally, the effect of Age on perception of the Cantonese contrasts for the Dutch controls was also investigated. As Table F.1.1d shows, Age was only correlated with perception of the Cantonese tone contrast.

Table F.1.1a

Kendall's Tau correlation analysis for the 22 Cantonese adoptees

	Accuracy Overall	Pretest	Intermediate test	Posttest	Affricates	Tones
Age	.149	.175	.196	.151	.107	.225
AoA	.057	.056	-.126	.085	.000	-.018
LoR	.118	.115	.200	.084	.084	.193

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.1.1b

Partial correlation analysis for the 22 Cantonese adoptees while controlling LoR

	Accuracy Overall	Pretest	Intermediate test	Posttest	Affricates	Tones
Age	.228	.297	.024	.212	.173	.233
AoA	.228	.297	.024	.212	.173	.233

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.1.1c

Descriptions of the two sub-groups of the Cantonese adoptees

Sub-group	No. of adoptees	Age	AoA	LoR
1	14	4;4-9;6 years	1;7-4;6 years	1;5-7;11 years
2	8	6;3-10;10 years	0;9-1;4 years	5;2-9;11 years

Table F.1.1d

Kendall's Tau correlation analysis for the two sub-groups of Cantonese adoptees and the Dutch controls

Sub-group 1	Accuracy Overall	Pretest	Intermediate test	Posttest	Affricates	Tones
Age	.500*	.366	.258	.596**	.457*	.433*
AoA	.034	.162	.045	.011	.046	-.079
LoR	.411*	.229	.236	.438*	.343	.367
Sub-group 2						
Age	-.357	.000	-.036	-.429	-.473	.036
AoA	-.113	-.245	-.077	-.189	-.269	.000
LoR	-.327	.039	.000	-.400	-.444	.074
Dutch controls						
Age	.171	.065	.257	.151	.012	.345*

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

F.1.2 Mandarin

For the Mandarin adoptees, LoR is significantly correlated with Age ($r = .867, p < .001$), and with AoA ($r = -.636, p < .001$). Firstly, Kendall's Tau correlation for the 26 Mandarin adoptees revealed several significant correlations for both Age and LoR (see Table F.1.2a). In specific, Age was reported to be positively correlated with overall perception of the Mandarin contrasts, and perception of the contrasts at the intermediate test, at the posttest, and perception particularly for the tone contrast. As for LoR, the results again partially overlapped those of Age, namely, perception of the Cantonese contrasts at the intermediate test, and perception particularly for the tone contrast. No other significant correlation was found.

Secondly, partial correlations for Age and AoA, while controlling LoR largely confirmed the results described above (particularly for Age). Age showed positive correlations with overall perception of the Mandarin contrasts, and perception of the contrasts at the posttest, and for both contrasts. Again identical results of Age and AoA were reported due to their linear relation.

Finally, following the same procedures in the Cantonese section, the 26 adoptees were re-grouped into two sub-groups (see Table F.1.2c), in which AoA and LoR were no longer significantly correlated anymore. Sub-group 1 consisted of 19 adoptees who were adopted between the age of 18 and 50 months (i.e., 1;6 - 4;2 years); sub-group 2 consisted of 6 adoptees who were adopted between the age of 10 and 16 months (i.e., 0;10 - 1;4 years). One adoptee was excluded from the re-grouping procedure due to his much older age of adoption (i.e., 68 months) compared to the other children, as described in chapter 3. The results of Kendall's Tau correlations for each sub-group are presented in Table F.1.2d. Only in the sub-group of 19 Mandarin adoptees, significant correlations were found, which confirmed those described in the first analysis. In particular, Age was positively correlated with overall perception of the Mandarin contrasts, and perception of the contrasts at the intermediate test, at the posttest, and perception particularly for the tone contrast, while LoR with perception of the contrasts at the intermediate test, and for the tone contrast (partially overlapping the results of Age). Most strikingly, a significant positive correlation between AoA and perception of the Mandarin contrasts at the pretest was found. As for the other sub-group of 6 adoptees, no significant correlation was found. Additionally, for the Dutch controls, Table F.1.2d shows no significant correlations with Age.

Table F.1.2a

Kendall's Tau correlation analysis for the 26 Mandarin adoptees

	Accuracy Overall	Pretest	Intermediate test	Posttest	Affricates	Tones
Age	.372*	.006	.387**	.364*	.185	.428**
AoA	.084	.179	-.051	.104	.047	.034
LoR	.246	-.054	.334*	.218	.073	.321*

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.1.2b

Partial correlation analysis for the 26 Mandarin adoptees while controlling LoR

	Accuracy Overall	Pretest	Intermediate test	Posttest	Affricates	Tones
Age	.547**	.226	.390	.505*	.462*	.414*
AoA	.547**	.226	.390	.505*	.462*	.414*

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

Table F.1.2c

Description of the sub-groups of Mandarin adoptees

Sub-group	No. of participants	Age	LoR	AoA
1	19	4;1-10;10 years	1;1-8;10 years	1;6-4;2 years
2	6	5;2-10;8 years	4;4-9;10 years	0;10-1;4 years

Table F.1.2d

Kendall's Tau correlation analysis for the two sub-groups of Mandarin adoptees and the Dutch controls

Sub-Group 1	Accuracy Overall	Pretest	Intermediate test	Posttest	Affricates	Tones
Age	.378*	-.061	.478**	.346*	.174	.441**
AoA	.089	.416*	-.150	.054	.042	.024
LoR	.327	-.129	.476**	.283	.127	.427*
Sub-Group 2						
Age	.467	.333	.358	.414	.467	.333
AoA	-.501	-.501	.154	-.296	-.501	-.358
LoR	.467	.333	.358	.414	.467	.333
Dutch controls						
Age	-.022	.076	-.174	.032	-.087	-.019

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

F. 2 Relationships between birth language production and Age, AoA, and LoR

In the present section, statistical analyses (Kendall's Tau correlation) for the effects of the three 'time variables' (Age, AoA, and LoR) on the production of the Chinese contrasts are presented. Again, Age, AoA, and LoR have a linear relation, in which the correlation between Age and LoR was inevitable. Similar to F.1, several performance measures were used: 1) Accuracy Overall, 2) at the Pretest, 3) at the Intermediate test, 4) at the Posttest, and 5) and 6) for each Target sound in each contrast (For Cantonese: Alveolar affricates [ts] and [ts^h], Tone 2 and Tone 5; For Mandarin: Retroflex affricates [tʂ] and [tʂ^h], Tone 2 and Tone 3). Following the same procedures as in F.1, three correlational analyses were conducted.

F.2.1 Cantonese

Firstly, Kendall's Tau correlations between each of the time variables and each of the performance measures for the complete group of Mandarin adoptees were carried out. As Tables F.2.1a, F.2.1b, F.2.1c, and F.2.1d show, only for the Cantonese affricate contrast, several significant correlations were reported, particularly for Age and LoR. Specifically, in the identification task, both Age and LoR were positively correlated with production of the Cantonese affricate contrast at the posttest. In the rating task, Age was positively correlated with overall production of the Cantonese affricate contrast, and production at the pretest and at the posttest, while LoR was positively correlated with production of the Cantonese affricate contrast at the posttest (partially overlapping the results of Age). No other significant correlations were found.

Secondly, partial correlations for Age and AoA while controlling LoR for all 26 adoptees were applied. As shown in Tables F.2.1e, F.2.1f, F.2.1g, and F.2.1h, there was no significant correlation for either Age or AoA in any task. Note that the outcomes for Age and AoA were identical due to their linear relation.

Finally, as in F.1.1, the 22 adoptees were assigned into two sub-groups (see Table F.1.1c), thus in either sub-group, Age and AoA were not correlated with each other anymore. Kendall's Tau correlations for each sub-group of adoptees in each task were carried out. Results in Tables F.2.1i, F.2.1j, F.2.1k, and F.2.1l show that only for the sub-group of 14 adoptees in the affricate rating task, Age and LoR were found to be significantly correlated with several measures of task performance, in particular, Age with overall production of the affricate contrast, and production at the pretest, at the intermediate test, at the posttest, and

production particularly for the aspirated affricate [ts^h]; while LoR with overall production of the affricate contrast, and production at the pretest, at the posttest, and production particularly for the unaspirated affricate [ts], largely overlapping with the results of Age. No other significant correlations were found for the adoptees. As for the Dutch controls, interestingly, several significant correlations with Age were found, but only for the tone contrast. In particular, in the tone identification task, Age was positively correlated with overall production of the tone contrast, and production at the posttest, and production particularly for Cantonese Tone 2. In the tone rating task, Age was positively correlated with production of the tone contrast at the intermediate test, at the posttest, and production particularly for Tone 2.

Table F.2.1a

Kendall's Tau correlation analysis for the 22 Cantonese adoptees in the affricate identification task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[ts]	[ts ^h]
Age	.193	.109	.128	.328*	.079	.079
AoA	.013	.264	-.177	-.121	.102	-.204
LoR	.144	-.061	.149	.332*	.040	.093

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1b

Kendall's Tau correlation analysis for the 22 Cantonese adoptees in the tone identification task

	Accuracy Overall	Pretest	Intermediate test	Posttest	Tone 2	Tone 5
Age	.000	-.190	.022	-.018	.053	-.188
AoA	.146	.187	.106	.093	-.057	.176
LoR	-.066	-.216	-.053	-.031	.048	-.183

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant the 0.01 level (2-tailed)

Table F.2.1c

Kendall's Tau correlation analysis for 22 Cantonese adoptees in the affricate rating task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[ts]	[ts ^h]
Age	.362*	.319*	.240	.385*	.258	.231
AoA	-.079	.000	-.237	-.022	.026	-.123
LoR	.279	.235	.279	.319*	.244	.209

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1d

Kendall's Tau correlation analysis for the 22 Cantonese adoptees in the tone rating task

	Accuracy Overall	Pretest	Intermediate test	Posttest	Tone 2	Tone 5
Age	.022	-.123	.053	.026	.144	-.193
AoA	.079	.190	-.031	.066	.097	-.057
LoR	-.017	-.136	.022	.022	.061	-.066

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1e

Partial correlation analysis for the 22 Cantonese adoptees in the affricate identification task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[ts]	[ts ^h]
Age	.169	.321	-.110	.158	.127	.073
AoA	.169	.321	-.110	.158	.127	.073

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1f

Partial correlation analysis for the 22 Cantonese adoptees in the tone identification task

	Accuracy Overall	Pretest	Intermediate test	Posttest	Tone 2	Tone 5
Age	.129	.070	.180	.070	.102	.048
AoA	.129	.070	.180	.070	.102	.048

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1g

Partial correlation analysis for the 22 Cantonese adoptees in the affricate rating task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[ts]	[ts ^h]
Age	.320	.354	.165	.361	.283	.193
AoA	.320	.354	.165	.361	.283	.193

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1h

Partial correlation analysis for the 22 Cantonese adoptees in the tone rating task

	Accuracy Overall	Pretest	Intermediate test	Posttest	Tone 2	Tone 5
Age	.081	-.014	.041	.180	.289	-.150
AoA	.081	-.014	.041	.180	.289	-.150

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1i

Kendall's Tau correlation analysis for the two sub-groups of Cantonese adoptees and the Dutch controls in the affricate identification task

Sub-group 1	Accuracy Overall	Pretest	Intermediate test	Posttest	[ts]	[ts ^h]
Age	.322	.278	.179	.375	.067	.134
AoA	-.056	.258	-.068	-.126	-.203	-.147
LoR	.233	.100	.089	.375	.089	.112
Sub-group 2						
Age	-.182	-.071	-.143	-.109	.071	.036
AoA	-.462	-.491	-.038	-.231	.265	-.231
LoR	-.148	-.036	-.109	-.074	.109	.000
Dutch controls						
Age	-.052	-.104	-.039	-.074	-.260	.096

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1j

Kendall's Tau correlation analysis the two sub-groups of Cantonese adoptees and the Dutch controls in the tone identification task

Sub-group 1	Accuracy Overall	Pretest	Intermediate test	Posttest	Tone 2	Tone 5
Age	-.179	-.316	-.078	-.156	-.078	-.167
AoA	.266	.023	.214	-.023	-.101	.281
LoR	-.201	-.192	-.167	-.112	-.056	-.233
Sub-group 2						
Age	-.182	-.143	-.214	.071	-.071	-.357
AoA	.038	.189	.113	.113	.038	.265
LoR	-.148	-.109	-.182	.109	-.036	-.327
Dutch controls						
Age	.356*	.162	.209	.314*	.416**	-.087

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.1k

Kendall's Tau correlation analysis the two sub-groups of Cantonese adoptees and the Dutch controls in the affricate rating task

Sub-group 1	Accuracy Overall	Pretest	Intermediate test	Posttest	[ts]	[ts ^h]
Age	.544**	.522**	.500*	.648**	.389	.456*
AoA	.034	.056	-.101	.045	-.101	.079
LoR	.433*	.411*	.389	.581**	.456*	.389
Sub-group 2						
Age	.286	.429	.214	.357	.214	.214
AoA	.113	.038	-.113	.265	.416	-.265
LoR	.255	.400	.182	.327	.255	.182
Dutch controls						
Age	.000	-.004	.013	.000	-.069	-.017

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.11

Kendall's Tau correlation analysis the two sub-groups of Cantonese adoptees and the Dutch controls in the tone rating task

Sub-group 1	Accuracy Overall	Pretest	Intermediate test	Posttest	Tone 2	Tone 5
Age	-.100	-.167	.011	.011	.189	-.211
AoA	-.034	.079	-.034	.258	.079	.011
LoR	-.011	-.033	-.056	.011	.100	-.100
Sub-group 2						
Age	.000	.000	.000	.000	.071	-.286
AoA	-.113	-.038	-.189	.038	-.340	-.189
LoR	.036	.036	.036	.036	.109	-.255
Dutch controls						
Age	.295	.013	.312*	.462**	.373*	.061

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

F.2.2 Mandarin

Firstly, Kendall's Tau correlations between each time variable and each performance measure for the 26 adoptees revealed several positive and negative correlations, as shown in Tables F.2.2a, F.2.2b, F.2.2c and F.2.2d. *Positive correlations:* Age and LoR with production of the Mandarin tone contrast at the posttest, in the rating task. *Negative correlations:* Age with production of the Mandarin affricate contrast at the pretest, and LoR with overall production of the Mandarin affricate contrast, production at the pretest, and production particularly for the unaspirated affricate [tʂ] in the identification task; LoR with production of the Mandarin tone contrast at the pretest and production for Tone 2 in the identification task. No significant correlations were found in the affricate rating task.

Secondly, partial correlations for the 26 adoptees between each performance measure and Age and AoA, while controlling LoR were carried out for each task. Tables F.2.2e, F.2.2f, F.2.2g, and F.2.2h, showed no significant correlations in any task. Again, the results for Age and AoA were identical due to their linear relation.

Finally, as in F.2.1, the 26 Mandarin adoptees were divided into two sub-groups (see Table F.1.2c), such that correlation between AoA and LoR was no longer significant in either

sub-group. Kendall's Tau correlations for each sub-group of adoptees in each task were conducted. As shown in Tables F.2.2i, F.2.2j, F.2.2k, and F.2.2l, only in the identification task, Age and LoR showed significant correlations with production of the Mandarin affricates at the pretest, and with production of Mandarin Tone 2. No other significant correlations for the adoptees were found. For the Dutch controls, (only) in the affricate rating task, Age showed a significant positive correlation with production of the unaspirated affricate [tʂ].

Table F.2.2a

Kendall's Tau correlation analysis for the 26 Mandarin adoptees in the affricate identification task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[tʂ]	[tʂ ^h]
Age	-.197	-.305*	-.016	.003	-.185	.040
AoA	.220	.116	.237	.068	.238	-.050
LoR	-.288*	-.291*	-.156	-.044	-.300*	.075

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2b

Kendall's Tau correlation analysis for the 26 Mandarin adoptees in the tone identification task

	Accuracy Overall	Pretest	Intermediate test	Posttest	Tone 2	Tone 3
Age	.084	-.241	.257	.140	-.210	.239
AoA	.118	.158	-.037	.175	.276*	.056
LoR	.000	-.289*	.199	.012	-.301*	.137

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2c

Kendall's Tau correlation analysis for the 26 Mandarin adoptees in the affricate rating task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[tʂ]	[tʂ ^h]
Age	-.052	-.111	.114	-.053	-.120	.059
AoA	-.043	-.053	-.019	-.019	.050	-.205
LoR	.006	-.016	.143	-.019	-.099	.149

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2d

Kendall's Tau correlation analysis for the 26 Mandarin adoptees in the tone rating task

	Accuracy Overall	Pretest	Intermediate test	Posttest	Tone 2	Tone 3
Age	.234	.046	.223	.316*	.146	.194
AoA	-.090	.012	-.121	-.097	.019	.012
LoR	.207	.000	.183	.277*	.118	.155

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2e

Partial correlation analysis for the 26 Mandarin adoptees in the affricate identification task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[tʂ]	[tʂ ^h]
Age	.100	-.207	.299	.055	.239	-.199
AoA	.100	-.207	.299	.055	.239	-.199

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2f

Partial correlation analysis for the 26 Mandarin adoptees in the tone identification task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[tʂ]	[tʂ ^h]
Age	.199	-.040	.153	.282	-.072	.303
AoA	.199	-.040	.153	.282	-.072	.303

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2g

Partial correlation analysis for the 26 Mandarin adoptees in the affricate rating task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[tʂ]	[tʂ ^h]
Age	-.336	-.395	-.130	-.330	-.170	-.345
AoA	-.336	-.395	-.130	-.330	-.170	-.345

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2h

Partial correlation analysis for the 26 Mandarin adoptees in the tone rating task

	Accuracy Overall	Pretest	Intermediate test	Posttest	[tʂ]	[tʂ ^h]
Age	.169	.032	.106	.139	-.166	.289
AoA	.169	.032	.106	.139	-.166	.289

Note: * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2i

Kendall's Tau correlation analysis for the two sub-groups of Mandarin adoptees and Dutch controls in the affricate identification test

Sub-group 1	Overall accuracy	Pretest	Intermediate test	Posttest	[tʂ]	[tʂ ^h]
Age	-.141	-.427*	.131	.112	-.117	.047
AoA	.076	.137	.196	.071	.088	.183
LoR	-.219	-.425*	.006	.024	-.172	-.030
Sub-group 2						
Age	-.467	.067	-.600	-.600	-.333	.138
AoA	-.072	-.501	.072	.072	-.215	.445
LoR	-.467	.067	-.600	-.600	-.333	.138
Dutch controls						
Age	-.056	-.003	-.168	-.050	.121	-.127

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2j

Kendall's Tau correlation analysis for the two sub-groups of Mandarin adoptees and Dutch controls in the tone identification test

Sub-group 1	Overall accuracy	Pretest	Intermediate test	Posttest	Tone 2	Tone 3
Age	.018	-.229	.177	.024	-.359*	.307
AoA	.106	.165	.006	.208	.307	-.065
LoR	-.059	-.297	.101	-.078	-.392*	.208
Sub-group 2						
Age	.333	-.067	.467	.600	.467	.333
AoA	.215	.215	.358	-.072	.072	.215
LoR	.333	-.067	.467	.600	.467	.333
Dutch controls						
Age	.206	.124	.066	.199	.031	.112

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.2k

Kendall's Tau correlation analysis for the two sub-groups of Mandarin adoptees and Dutch controls in the affricate rating test

Sub-group 1	Overall accuracy	Pretest	Intermediate test	Posttest	[tʂ]	[tʂʰ]
Age	-.047	-.241	.265	.012	-.182	.101
AoA	.018	.024	-.106	.065	-.024	-.024
LoR	-.041	-.166	.249	.006	-.142	.084
Sub-group 2						
Age	.067	.200	-.333	-.600	-.200	-.200
AoA	-.501	-.501	-.072	-.215	-.501	.072
LoR	.067	.200	-.333	-.600	-.200	-.200
Dutch controls						
Age	.158	.174	.081	.081	.304*	-.068

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

Table F.2.21

Kendall's Tau correlation analysis for the two sub-groups of Mandarin adoptees and Dutch controls in the tone rating test

Sub-group 1	Overall accuracy	Pretest	Intermediate test	Posttest	Tone 2	Tone 3
Age	.188	.088	.124	.307	.142	.253
AoA	-.159	-.047	-.165	.231	.018	-.142
LoR	.183	.024	.119	.280	.089	.226
Sub-group 2						
Age	.200	-.067	.067	.467	.200	.067
AoA	.358	.215	.358	.358	.072	.645
LoR	.200	-.067	.067	.467	.200	.067
Dutch controls						
Age	.149	.050	.162	.193	.127	.078

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

F.3 Relationships among residual memory of birth language vocabulary and Age, AoA, and LoR

This section presents the statistical analyses (Kendall's Tau correlation) for the effects of the three 'time variables' (Age, AoA, and LoR) on the residual memory of birth language vocabulary. Four measures of task performance were used, i.e., Accuracy Overall, at the Pretest, at the Intermediate test, and at the Posttest. The same three correlational analyses were used.

F.3.1 Cantonese

Firstly, Kendall's Tau correlations between each time variable and each performance measure for the 22 adoptees were applied. As Table F.3.1a shows, there was no significant correlation with any time variable.

Secondly, Partial correlations for Age and AoA, while controlling for LoR confirmed the results described in the previous analysis. There was no significant correlation, as presented in Table F.3.1b. Note again that the outcomes of the correlations for Age and AoA were identical because of their linear relation.

Finally, again the 22 Cantonese adoptees were divided into two sub-groups (see Table F.1.1c), thus the significant correlation between AoA and LoR were disentangled in both sub-groups. Kendall's Tau correlations revealed several significant correlations particularly in the sub-group of 14 adoptees. Age had a positive correlation with Cantonese vocabulary identification at the posttest; LoR had positive correlations with overall accuracy of Cantonese vocabulary identification and at the pretest (partially overlapping with the results of Age). Interestingly, AoA showed a negative correlation with Cantonese vocabulary identification at the pretest. Within the group of Dutch controls, Age also showed a positive correlation with overall accuracy of Cantonese vocabulary identification.

Table F.3.1a

Kendall's Tau correlation analysis for the 22 Cantonese adoptees

	Accuracy Overall	Pretest	Intermediate test	Posttest
Age	.145	.161	.078	.140
AoA	.128	-.131	.202	.000
LoR	.105	.201	.027	.113

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

Table F.3.1b

Partial correlation analysis for the 22 Cantonese adoptees

	Accuracy Overall	Pretest	Intermediate test	Posttest
Age	.177	-.009	.154	.106
AoA	.177	-.009	.154	.106

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

Table F.3.1c

Kendall's Tau correlation analysis for the two sub-groups of Cantonese adoptees and the Dutch controls

Sub-group 1	Accuracy Overall	Pretest	Intermediate test	Posttest
Age	.313	.301	.023	.441*
LoR	.402*	.416*	.069	.215
AoA	-.283	-.421*	.000	.036
Sub-group 2				
Age	.036	.036	.491	-.214
LoR	.000	.000	.539	-.182
AoA	.539	.000	.240	.038
Dutch controls				
Age	.298*	.126	.271	.215

Note. * Correlation is significant at the 0.05 level (2-tailed)

** Correlation is significant at the 0.01 level (2-tailed)

F.3.2 Mandarin

Firstly, Kendall's Tau correlations between each of the time variables and each of the six performance measures for the 26 adoptees showed no significant correlations, as shown in Table F.3.2a.

Secondly, Partial correlations for Age and AoA, while controlling for LoR confirmed previous results, i.e., there was no significant correlation with any time variable, as shown in Table F.3.2b. (Note that same results of Age and AoA were again reported due to their linear relation.)

Finally, the 26 Mandarin adoptees were re-grouped into two sub-groups, in which AoA and LoR were no longer correlated in either group. Kendall's Tau correlations showed Age and LoR were negatively correlated with Mandarin vocabulary identification at the posttest, only for the six adoptees who were adopted between 10 and 16 months old. No other significant correlations were found.

Table F.3.2a

Kendall's Tau correlation analysis for the 26 Mandarin adoptees

	Accuracy Overall	Pretest	Intermediate test	Posttest
Age	-.201	-.080	-.158	-.169
AoA	-.043	.186	-.201	.067
LoR	-.136	-.125	-.058	-.150

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.3.2b

Partial correlation analysis for the 26 Mandarin adoptees

	Accuracy Overall	Pretest	Intermediate test	Posttest
Age	.123	.295	-.165	.092
AoA	.123	.295	-.165	.092

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Table F.3.2c

Kendall's Tau correlation analysis for the two sub-groups of Mandarin adoptees and the Dutch controls

Sub-group 1	Overall accuracy	Pretest	Intermediate test	Posttest
Age	-.246	-.108	-.255	-.103
LoR	-.195	-.097	-.226	-.067
AoA	.018	.139	-.100	.042
Sub-group 2				
Age	.067	.358	.067	-.828*
LoR	.067	.358	.067	-.828*
AoA	.645	.077	.358	.000
Dutch controls				
Age	-.047	-.013	-.029	-.026

Note. * Correlation is significant at the 0.05 level (2-tailed)
 ** Correlation is significant at the 0.01 level (2-tailed)

Nederlandse samenvatting

Internationaal geadopteerde kinderen vertonen een uniek patroon in hun taalontwikkeling dat sterk verschilt van dat van niet-geadopteerde (éentalige of tweetalige) kinderen. Terwijl adoptiekinderen vóór de adoptie (dus gedurende een relatief korte periode) exclusief met de taal van het land van herkomst in aanraking komen (die ik vanaf nu hun ‘oorspronkelijke taal’ zal noemen), worden zij door hun adoptie volledig van die taal afgesneden, om vervolgens exclusieve input te ontvangen in de taal van hun nieuwe land. Er werd vaak aangenomen dat adoptiekinderen hun ‘oorspronkelijke taal’ na de adoptie vergeten. In dit onderzoek zijn twee groepen jonge Chinese kinderen onderzocht die zijn geadopteerd door Nederlandssprekende gezinnen in Nederland, één groep uit delen van China waar Mandarijn Chinees wordt gesproken, en één groep uit delen van China waar Cantonees Chinees wordt gesproken. Ten tijde van het onderzoek waren deze kinderen tussen vier en tien jaar oud, en hadden ze gedurende gemiddeld vijf jaar geen contact meer gehad met hun oorspronkelijke taal. Dit onderzoek richt zich op drie vragen: a) Zijn jonge geadopteerde Chinese kinderen enkele jaren na de adoptie nog in staat om fonologische contrasten uit hun oorspronkelijke taal (namelijk affricaten en tonen) te herkennen (verstaan) en te produceren (uit te spreken)? b) Hebben ze nog bewuste herinneringen aan het vocabulaire van de oorspronkelijke taal? c) Als het antwoord op bovenstaande vragen ‘ja’ is, zijn deze herinneringen dan onmiddellijk beschikbaar, of kunnen ze slechts toegankelijk gemaakt worden door hernieuwde blootstelling aan de oorspronkelijke taal?

In sommige voorgaande studies werden er bij volwassenen die geadopteerd waren, nadat ze enige tientallen jaren geen contact hadden gehad met de oorspronkelijke taal, helemaal geen herinneringen meer gevonden aan de oorspronkelijke taal (Pallier et al., 2003; Ventureyra, Pallier, & Yoo, 2004). Uit andere studies bleek echter dat geadopteerden in de tienerleeftijd en op volwassen leeftijd door hernieuwde blootstelling aan de oorspronkelijke taal hun kennis van de fonologie van deze taal konden herwinnen (Choi, 2014; Singh et al., 2011; Pierce et al., 2014). Hoofdstuk 3 van dit proefschrift onderzocht de herkenning van Chinese fonologische contrasten (namelijk affricaat- en tooncontrasten) voor en na perceptuele training. Het beschrijft een experiment met Mandarijn Chinees en een experiment met Cantonees Chinees. Beide experimenten laten zien dat Chinese adoptiekinderen de contrasten uit hun moedertaal beter herkenden dan niet-geadopteerde Nederlandse kinderen (die met de adoptiekinderen gematcht waren op verschillende factoren), *nadat* ze de

perceptuele training hadden voltooid. Voor en tijdens de training verschilden de geadopteerde en niet-geadopteerde kinderen niet van elkaar in hun herkenning van de fonologische contrasten. Verder bleek een betere herkenning soms samen te hangen met een hogere adoptieleeftijd, of met contact met de oorspronkelijke taal na de adoptie (door de aanwezigheid van geadopteerde broertjes of zusjes uit hetzelfde taalgebied, of door bezoeken aan China). Dit onderzoek versterkt de conclusies uit voorgaand onderzoek dat de ervaring die kinderen opdoen met hun oorspronkelijke taal, al is het maar gedurende een korte periode, toch leidt tot fonologische kennis die goed bewaard blijft, en met hernieuwde blootstelling aan de oorspronkelijke taal weer toegankelijk kan worden. Tegelijk bieden de resultaten uit dit onderzoek ook voor het eerst empirisch bewijs dat geadopteerden hun moedertaal al snel na de adoptie ‘vergeten’.

Een interessante vergelijking kan worden gemaakt met mensen die opgroeiden met de meerderheidstaal van hun land, maar van wie de ouders of grootouders (soms) ook een minderheidstaal gebruikten (die ik vanaf nu hun ‘erftaal’ zal noemen). Sommigen van deze mensen werden als kind eerst regelmatig blootgesteld aan de erftaal, maar vanaf een bepaald moment bijna niet meer. Wanneer zij, eenmaal volwassen, de erftaal weer gingen leren bleken zij beter te zijn in het leren van de uitspraak van deze taal dan mensen die als kind geen ervaring hadden opgedaan in deze taal (Au et al., 2002; Oh et al., 2003; Knightly et al., 2003; Au et al., 2008). Een belangrijk verschil tussen deze mensen met een erftaal en geadopteerden is dat de eersten altijd een bepaalde mate van contact met de taal behielden, in tegenstelling tot de geadopteerden. Er is slechts heel weinig onderzoek gedaan naar het leren van de uitspraak van de oorspronkelijke taal door geadopteerden; de enige studie op dit gebied is een recent proefschrift van Choi (2014). Choi (2014) onderzocht volwassenen die als jonge kinderen waren geadopteerd uit Korea en liet zien dat zij bij hernieuwde blootstelling aan het Koreaans de uitspraak van Koreaanse klanken beter onder de knie kregen dan een niet-geadopteerde controlegroep. Deze resultaten waren dus vergelijkbaar met die van de erftaal-leerders. Hoofdstuk 4 van dit proefschrift onderzocht de uitspraak van Chinese fonologische contrasten door de jonge Chinese geadopteerden voor en na perceptuele training. Evenals hoofdstuk 3 beschrijft ook hoofdstuk 4 een experiment met Mandarijn Chinees en een experiment met Cantonees Chinees. In beide experimenten werd een robuust verschil gevonden tussen de uitspraak van de geadopteerde en de niet-geadopteerde kinderen. De uitspraak van de geadopteerde kinderen was consequent beter dan die van de niet-geadopteerde controlegroep. Interessant genoeg was dit verschil niet alleen

aanwezig na de perceptuele training, maar ook daarvoor. (In tegenstelling tot hoofdstuk 3 werd er hier geen duidelijk effect gevonden van adoptieleeftijd en contact met de oorspronkelijke taal na de adoptie.) De resultaten van dit proefschrift stemmen dus overeen met de bevindingen uit eerder onderzoek bij *volwassen* erftaal-leerders en Koreaanse geadopteerden; evenals deze voorgaande studies laat dit proefschrift zien dat fonologische kennis die is verworven in de kindertijd een langdurig effect kan hebben niet alleen op de perceptie maar ook op de productie van klanken. Opmerkelijk is dat het verschil tussen geadopteerden en niet-geadopteerden in dit proefschrift al optrad voor de perceptuele training, en in het onderzoek van Choi (2014) bij volwassen geadopteerden pas na de perceptuele training, wat wijst op het belang van de tijd die is verstreken sinds de adoptie.

Bij internationale adoptie wordt de communicatieve functie van de oorspronkelijke taal overgenomen door de taal van de nieuwe omgeving, wat er toe zou kunnen leiden dat geadopteerde kinderen het vocabulaire van de oorspronkelijke taal vergeten. Isurin (2010) beschrijft een Russisch meisje dat op negenjarige leeftijd geadopteerd werd en geleidelijk haar productieve en perceptuele Russische vocabulaire verloor. Andere studies (met grotere aantallen proefpersonen) laten zien dat er bij geadopteerden, eenmaal volwassen, geen herinneringen meer worden gevonden aan woorden uit de oorspronkelijke taal (Oh et al., 2010; Choi, 2014). Schmid (2002) beschrijft daarentegen dat Duits-Joodse emigranten, waaronder ook kinderen, hun vocabulaire in de moedertaal vele decennia na hun emigratie behielden, ondanks zeer beperkte input in die taal na de emigratie. Hoofdstuk 5 van dit proefschrift onderzocht of de jonge Chinese adoptiekinderen nog bewuste herinneringen bewaarden aan het vocabulaire van hun oorspronkelijke taal, wederom veel korter na de adoptie dan de meeste voorgaande studies. De resultaten bevestigden die van de voorgaande studies en lieten geen herinneringen zien aan het vocabulaire van de oorspronkelijke taal bij de Mandarijnse en Cantonese geadopteerde kinderen, ook niet na de perceptuele training. Het verschil tussen deze uitkomst en de bevindingen van Schmid (2002) lijkt te wijzen op het belang van de hoeveel input in de oorspronkelijke taal. Net als in hoofdstuk 4 was er in hoofdstuk 5 geen sterk effect van adoptieleeftijd en contact met de oorspronkelijke taal na de adoptie door geadopteerde broertjes en zusjes en bezoeken aan China. Voor de Mandarijnse geadopteerde kinderen bleek een groter vocabulaire in het Nederlands samen te hangen met een kleiner vocabulaire in het Mandarijn Chinees. De resultaten uit hoofdstuk 5 bieden verdere evidentie voor de conclusie dat de kennis van de oorspronkelijke taal achteruit gaat wanneer het contact met deze taal wordt verbroken.

Kortom, dit is het eerste onderzoek dat empirisch bewijs heeft geleverd, op grond van een groot aantal jonge Chinese adoptiekinderen, dat herinneringen aan de oorspronkelijk taal behouden blijven. De onderzoeksresultaten lieten zien dat de geadopteerde kinderen aanvankelijk geen blijk gaven van enige kennis van de klanken van hun oorspronkelijke taal, maar wel na perceptuele training van deze klanken. De geadopteerde kinderen bewaarden geen meetbare herinneringen aan het vocabulaire van hun oorspronkelijke taal. Verder biedt dit onderzoek, ondanks de beperkingen die een relatief kleine testgroep biedt voor correlatie-analyse, inzichten in het effect van de adoptieleeftijd en van contact met de oorspronkelijke taal na de adoptie. Tenslotte toont dit onderzoek dat de hoeveelheid tijd die is verstreken sinds de adoptie een belangrijke rol speelt bij de herinneringen aan de oorspronkelijke taal, met name waar het de uitspraak betreft.

中文概述

相对于单语和双语学习者，被跨国领养儿童的语言发展模式具有独特特点。在被领养前，这些儿童完全生活在母语环境下。在被领养后，他们的母语被收养他们的家庭的语言取而代之，因此被领养儿童的母语经验仅限于童年很短的一段时间内。人们普遍认为，孤儿被跨国收养后会很快忘记自己的母语。本博士课题调查了在中国粤语和普通话区域出生并被荷兰家庭领养几年后的中国儿童的母语语言能力。实验所考察儿童的年龄在四到十岁之间，他们与自己的母语隔离时间平均为五年。本课题的核心研究问题是：被跨国领养多年后的中国儿童能否感知(perception)和产生(production)其母语语音中一些对立对(特别是塞擦音和声调)。另外，被领养多年后他们是否还保留着对母语词汇知识。如果答案是肯定的，他们关于母语的记忆是否可以即刻呈现，还是需要重新接受母语输入后才能被激活呢？

一些研究发现，被跨国领养几十年后的韩国孤儿在没有任何母语重新输入的情况下已经没有了任何母语语音记忆的痕迹(Pallier et al., 2003; Ventureyra, Pallier, & Yoo, 2004)。然而，另外一些研究发现，被跨国领养的孩子即使进入青少年或成人时期，如果重新接受母语输入，他们也可以成功恢复母语语音知识(Choi, 2014; Singh et al., 2011; Pierce et al., 2014)。本篇论文第三章考察了被领养的中国儿童在语音训练之前和之后的母语语音(特别是塞擦音和声调)感知(perception)情况。研究发现，在接受语音培训前和语音培训过程中，被收养中国儿童的母语语音感知(perception)能力与那些实验前从来没有接触过任何中国语言的荷兰儿童并没有很大区别。然而在语音培训结束后，粤语和普通话两个实验的结果一致表明，被领养的中国儿童的母语语音感知(perception)能力显著地超过了荷兰儿童。另外，研究发现被领养年龄和重新接触母语语音似乎对母语语音知识保留和恢复有积极作用。具体来讲，在被领养时年龄较大的儿童会较好的保留母语语音知识，即使在被领养后没有任何母语重新输入(虽然此结果仅显示在普通话实验中)。通过偶然从被领养到同一家庭的中国儿童那里听到母语或者通过参观中国听到母语有助于母语语音感知(perception)记忆的恢复。总而言之，本研究结果首先为一直以来人们认为孤儿被跨国领养后会没有母语知识输入后会很快遗忘母语的这个观念提供了科学依据，另外本研究结果进一步支持了已有研究成果，说明 -

母语重新输入可以激活已有的母语语音记忆。更重要的是，本研究填补了被领养儿童母语语音记忆与恢复 - 在行为学研究领域 - 的空白。

过去研究显示，继承了童年时期掌握的语言的人(heritage language learners)，与该语言初学者相比，成年时期重新学习该语言后在语音产生(production)方面更胜一筹(Au et al., 2002; Oh et al., 2003; Knightly et al., 2003; Au et al., 2008)。值得一提的是，那些继承童年语言者在童年早期学习该语言后并没有完全断绝与该语言的联系，而是持续接触该语言(虽然其接触很有限)。与之相反，孤儿被跨国领养后完全失去了与自己母语的联系。目前，对被跨国领养人的母语语音产生(production)的研究非常有限。Choi (2014) 考察了被荷兰家庭领养几十年后的韩国孤儿的母语语音产生(production)情况，研究结果与之前语言继承者的调查结果相似。即在重新学习童年时期的语言后，韩国孤儿的母语语音产生(production)显著地超过了韩语初学者。然而，在重新学习之前，两组被试的成绩并没有显著区别。本论文第四章系统调查了被荷兰家庭领养的中国儿童在语音感知培训之前和之后的母语语音产生(production)情况。粤语和普通话两个实验一致表明被跨国领养的中国儿童在语音感知训练之前和之后在母语语音产生方面都有很大优势。此结果虽与之前关于成人时期的被收养的韩国人的实验结果相似，但是进一步丰富了过去的研究发现，即呈现了与母语隔绝时间相对比较短的中国儿童在母语重新输入前显示的母语语音产生(production)优势。这一成果凸显了与母语隔绝的时间长短对母语语音产生(production)的影响。与第三章感知研究不同，本章节中被领养儿童的被领养年龄和母语再渗入(通过偶然从被领养到同一家庭的中国儿童那里听到母语或者通过参观中国听到母语)对于语音产生(production)影响相对有限。这可能与本研究中的样本量对于相关分析较小有关。总之，本章的研究结果证明童年时期掌握的母语语音知识不仅对母语语音感知(perception)而且对母语语音产生(production)能力有长久影响。

当母语被领养语言取代以后，被跨国领养的儿童很可能会逐渐忘记母语的词汇知识。一项个案研究显示一位九岁的俄罗斯女孩在被领养后渐渐失去了母语词汇感知和产生能力(Isurin, 2000)。另外一些研究调查了更多的被领养人，发现这些被领养人在成年以后丧失了母语词汇知识的记忆(Oh et al., 2010; Choi, 2014)。然而，二战期间移居海外的德国犹太人在移民几十年后依旧保存了德语词汇知识(Schmid, 2002)，这可能是由于这些犹太移民在移居海外时比较年长(最小的年龄是 11 岁)，母语知识已经掌握

很成熟。本论文第五章探索了被领养的中国儿童在被领养几年后(与上述研究相比经历了更短的母语隔绝时间)是否还保存母语词汇记忆。研究结果表明,即使在语音感知培训之后,被收养中国儿童识别的母语词汇数量也没显著超越荷兰本土儿童。与 Schmid (2002)相比,本研究中中国儿童在被领养前接受母语知识的时间相对比较短,另外在被领养后完全失去了母语接触。这些却别似乎强调母语知识输入量(包括被领养之前和之后)的重要性。与第四章相似,被领养年龄和母语再输入(通过偶然从被领养到同一家庭的中国儿童那里听到母语或者通过参观中国听到母语)对于词汇记忆没有显著影响。值得一提的是,本章研究结果(尤其是普通话实验的结果)发现母语词汇知识保留和被领养语言发展之间存在负相关,即母语词汇记忆越好,其被领养语言能力越差。总之,本章研究结果支持了过去的研究结果,即母语输入被切断后母语知识会退化。

综上所述,本研究结果表明,与母语隔离几年之后,被跨国收养儿童的母语语音感知(perception)能力会变弱,母语词汇知识也会遗忘;但是,当重新接受母语知识(特别是语音知识)输入以后,这些儿童可以成功恢复母语的语音知识。值得一提的是,在语音产生(perception)方面,这些儿童在重新接受母语输入前就展示了超强的优势。另外,虽然本课题中被调查的儿童数量对于相关分析有一定的局限性,本研究在领养年龄和母语再渗入对母语知识保存的影响方面具有参考价值。最后,本课题是首个对大量被跨国领养几年后的儿童进行的关于母语知识记忆的系统调查。该研究填补了本领域研究在儿童群体和行为学研究领域的空白。

Curriculum Vitae

Wencui Zhou received her education both in China and in the Netherlands. In 2004, she gained her bachelor's diploma in English Education from Qiqihar University in China. Since then, she worked as a college lecturer in English for five years in China, first at Jiujiang University between 2004 and 2007, and later at HKU SPACE Global College Suzhou until 2009. Since 2009, she pursued her education in the Netherlands. She studied Applied Linguistics for a master's degree at the University of Groningen. Before finishing her Master's program, she was awarded a PhD scholarship at the Max Planck Institute for Psycholinguistics in Nijmegen (MPI). In the summer of 2010, immediately after receiving her master's degree, she started working as a PhD student in the Language Comprehension Department at the MPI. In 2013, she spent four months in China to collect data from native Chinese children and adults, which was funded by the MPI and the Internationalisation Fund from Radboud University. Currently, she is working as a research assistant for a joint project between University of Kansas (Department of Linguistics) and Radboud University (Center for Language Studies), while in the meantime pursuing opportunities to further advance her academic career.

Publications

Zhou, W., Broersma, M. & Cutler, A. (in preparation). Strength of birth language perception in young adoptees.

Zhou, W., Broersma, M. & Cutler, A. (in preparation). Resilient memories of young adoptees' birth language production.

Zhou, W., & Broersma, M. (2014). Perception of birth language tone contrast by adopted Chinese children. *In Proceedings of TAL 2014: The 4th International Symposium on Tonal Aspects of Languages*

Gussenhoven, C., & Zhou, W. (2013). Revisiting pitch slope and height effects on perceived duration. *In Proceedings of INTERSPEECH 2013: 14th Annual Conference of the International Speech Communication Association (pp. 1365-1369)*

MPI series in Psycholinguistics

1. The electrophysiology of speaking: investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turenhout*
2. The role of the syllable in speech production: evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: a self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: an Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: a study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: an experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: the case of phonetic sequences. *Andrea Christine Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: the selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*

22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: computational and cross-linguistic experimental studies. *Fermin Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: an electro-physiological approach. *Danielle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: the expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: an electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disuency: interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*

46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: an information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: evidence from electroencephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: an interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: a neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: extrinsic normalization in vowel perception. *Matthias J. Sjerps*

69. Structuring language: contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: a second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Connie de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*
85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*

90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants' understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sumer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. Incrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees. *Wencui Zhou*