The role of prosodic input in word learning

A cross-linguistic investigation of Dutch and Mandarin Chinese infant-directed speech

Published by
LOT
Kloveniersburgwal 48
1012 CX Amsterdam
The Netherlands

phone: +31 20 525 2461

e-mail: lot@uva.nl
http://www.lotschool.nl

Cover illustration: Longxiang Wang

ISBN: 978-94-6093-319-6
NUR 616

The role of prosodic input in word learning

A cross-linguistic investigation of Dutch and Mandarin Chinese infant-directed speech

De rol van prosodische input bij het leren van woorden:
Een cross-linguïstisch onderzoek naar kindgerichte spraak in het Nederlands en het Mandarijn
(met een samenvatting in het Nederlands)

Proefschrift
ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus,
prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op maandag 8 april 2019 des middags te 4.15 uur

door

Mengru Han
geboren op 10 juli 1989 te Yichang, China

Promotor: Prof. dr. R.W.J. Kager

Copromotor: Dr. N.H. de Jong

The research reported in this dissertation was supported by the China Scholarship Council and the Utrecht Institute of Linguistics (OTS).

Table of Contents

List of Tables

Chapter 6

List of Figures

Acknowledgments

It's been more than five years since I started my Ph.D. This journey has been challenging yet fruitful, and luckily, I have not undertaken it alone. I wish to thank those who have supported me along this journey and made this dissertation possible.

My deepest gratitude goes to my supervisors René Kager and Nivja de Jong. Really, I couldn't wish for better supervisors. René, your passion and dedication to research as well as your kindness and patience have influenced me deeply. I truly appreciate that you have been supportive and open-minded about my research. Thank you for gently guiding me towards becoming an independent researcher and thank you for always believing in me.

Nivja, I have learned so much from you, from Praat scripting to R, statistics, and more. I have greatly enjoyed our meetings (with scones) at the old de Bakkerswinkel at the Hague. I vividly remember the first time you showed me how to do statistics with R and encouraged me to try it out. You told me it would be fun, and it was! Thank you so much for your precision, enormous patience, and calmness.

I really appreciate those who agreed to be in my reading committee, including Barbara Höhle, Claartje Levelt, Nan Bernstein Ratner, Leher Singh, and Frank Wijnen. Thank you all for spending time reading my dissertation and giving me invaluable comments and suggestions.

I would like to thank many of the researchers who taught courses, offered their insights on my research, or otherwise helped me in one way or another: Huub van den Bergh, Hans Rutger Bosker, Titia Benders, Denis Burnham, Aoju Chen, Yiya Chen, Lisa Cheng, Alex Cristia, Carlos Gussenhoven, Vincent van Heuven, Chantal Kemner, Hugo Quené, Leher Singh, Henriette de Swart, and Hans van de Velde.

I would also like to thank Maaike Schoorlemmer for giving me guidance on settling in Utrecht and for always promptly answering my administrative questions. Brigitte Hertz, I benefited a lot from your project management course. Thank you for letting me know that I did well. Simone Veld, thanks for helping me adjust to life after I returned to work following maternity leave.

I am grateful to my teachers at Nanjing University: Xianwen Cao, Qinghua Ma, Xueli Yang, and Quanzhen Zhang, who introduced me to linguistics. Without your encouragement, I would not even have had the courage to study abroad. I also thank my teachers during my exchange in National University of Singapore: Ho Chee Lick, Lee Cher Leng, Adrian Tien, and Zheng Xu. That semester changed my life—I made up my mind to pursue my interest in linguistics after taking your courses. Special thanks go to Zheng Xu, who gave me his best wishes before I came to Utrecht.

My paranymphs have been giving me genuine support even before I started my Ph.D. Jeroen (Emoe) Breteler, thank you for keeping me up to date with the latest news in metrical phonology, for helping me to sharpen my views, for commenting on two chapters of this dissertation, and for translating my Dutch summary. I'm grateful that you kept Ally busy (or let Ally keep you busy) so I could work. Aside from work, thank you for countless lunches, for good coffee, rooibos tea, and Roberto Gelato, for offering your parenting insights, and your meditation and productivity tips. Sparkling Yan Gu, thanks for sharing your experiences with scholarship applications, for sending me new articles (at late nights or on your way to the airport!), and for all the chats, encouragement, and discussions. I'm glad that we can share all our excitements and disappointments together.

I wish to thank the Baby Lab members, with whom I have shared many Friday mornings in the basement of Janskerkhof. Frank Wijnen, thank you for your wisdom and humor. Maartje de Klerk, thank you for helping me set up my experiments and offering your genuine advice on my experiments. Ao Chen and Liquan Liu, thank you for setting good examples. Thanks also go to Tessel Boerma, Emma Everaert, Ileana Grama, Carolien van den Hazelkamp, Caroline Junge, Brigitta Keij, Annemarie Kerkhoff, and Rianne van Rooijen.

I am grateful to Run Chai, Suzanne Dekker, Willemijn Doedens, Lisanne Geurts, Karlijn Kouwer, Sule Kurtcebe, Yuchen Li, Judith van der Linden, and Cora Pots, for their help with coordinating experiments, data collection, and annotation. I am grateful to Aihua Zou and Yuhong Li from Taohualing Kindergarten, Huan He from Gezhouba Dongshan Kindergarten, and Hong Xie at Gezhouba Early Education Center in Yichang, China for their kind support and coordination in recruiting participants. I would also like to express my gratitude to all the families who participated in my experiments at the Utrecht Baby Lab and in China.

I also thank the UiL OTS Lab support: Maarten Duijndam, Jacco van Elst, Martijn van der Klis, and Chris van Run. You have made every researcher's life so much easier.

Trans was a maze to me at first, but after almost eight years I now very much consider it home—a sweet home with good memories. Thank you, my wonderful colleagues, for making the difference. Anqi Yang, we've had laughs together. Stay wild. Thank you, Xin Li, for sharing the final stages with me and helping me deliver my manuscripts. Xiaoli Dong, Ximeng Dong, and Bert le Bruyn, thanks for the lovely dinners and playdates. I would also like to extend my warm thanks to Frans Adriaans, Stavroula Alexandropoulou, Myrthe Bergstra, Andrea Santana Covarrubias, Quy Ngoc Thi Doan, Marjolein van Egmond, Shanshan Fan, Anja Goldschmidt, Haoyan

Ge, Björn't Hart, Marlisa Hommel, Luying Hou, Na Hu, Shuangshuang Hu, Bambang Kartono, Loes Koring, Martijn van der Klis, Heidi Klockmann, Tom Lentz, Fang Li, Yingtong Liu, Zenghui Liu, Sophia Manika, Chou Mo, Sandrien van Ommen, Liv Persson, Anne-France Pinget, Erin Pretorius, Silvia Rădulescu, Anna Sara Romøren, Alexia Guerra Rivera, Nina Sangers, Maartje Schulpen, Yuning Sun, Silvia Terenghi, Assaf Toledo, Jolien Scholten, Yuan Xie, Franca Wesseling, Yipu Wei, Jingwei Zhang, and Jun Zhang.

I would also like to acknowledge some of the other Chinese linguists I met in the Netherlands: Yifei Bi, Min Liu, Menghui Shi, He Sun, Jiang Wu, Junru Wu, Yang Yang, and Ting Zou.

Additional thanks go to my friends from high school and college, who have always been there for me despite time differences. Yiting Liu, Danni Shi, and Jingwen Zhu, you are the best. Xijinyan Chen, thank you for meeting me in London, hosting me in New York, and sharing your own experiences with the crazy journey of writing a dissertation with a toddler.

I would like to express my gratitude to The Pilot Family. Lucille, Arjen, Niño, and Louis, thank you for taking care of Ally and always being considerate to us. I also wish to thank the roommates I've had at one point or another in the past, namely Qingyi Feng, Miaoying Yang, and Liuchi Yi, for keeping me sane after work.

Special thanks to Sara, for proofreading most of my chapters and your suggestions on how to write.

I have received enormous support from my family in China. 感谢我的家人。我的爷爷奶奶们，感谢你们对我的爱和鼓励，我在荷兰常常想念你们。感谢公公、婆婆、小姨、小姨爹、姨妈、姨爹、姑妈、姑爹、表妹、表哥们和表嫂们一直以来对我的支持。

My dearest Ally, thank you for making me laugh and sharing your thoughts with me. You have tried your best to understand that mommy is always working. My husband Peng, thank you for always having faith in me. You are the most optimistic person I've ever known. I am truly grateful for your love, patience, and support.

I dedicate this dissertation to my parents. 爸爸，我感谢您对我无微不至的照顾、对我异乡求学的支持、以及对我和我女儿的无条件的爱。妈妈，我感谢您教会我对学业和事业的责任感以及对完美的追求，也感谢您数次放下工作去荷兰帮我照顾女儿。我将这本论文献给你们。

Kuala Lumpur,
February 2019

Chapter 1

Introduction

"Look! That's an excavator! See the excavator?" A mom points to an excavator, smiling, telling her two-year-old daughter using a high-pitched and exaggerated tone. This is the first time the girl has seen an excavator. The big, yellow construction vehicle with a big arm soon attracts her attention. She points to it, trying to repeat the word "excavator" after her mom. "Yes! An excavator!" her mom replies. The next time the girl sees an excavator, she probably will get excited and say "Excavator!"

This is a typical scene in which children learn words by interacting with their caregivers. As simple as such word-learning contexts may appear, learning words is, in fact, a formidable task from a linguistic point of view. In order to learn words, children need to be familiar with the sounds in their native language(s). They must be able to segment words from continuous speech, as words are usually embedded in sentences with no silent pauses between them. Importantly, they need to be sensitive to novel word labels in linguistic input and associate them with concepts, known as "word-to-world mapping" (Brooks & Kempe, 2014, p. 679). In the example illustrated above, the two-year-old girl should have already been familiar with the consonants and vowels in the sound form [ˈɛkskəˌveɪtər], and she has made considerable progress in segmenting words from continuous speech at this age (Jusczyk & Aslin, 1995). She is alert when she hears the word "excavator," as she has never heard of this word before and therefore the word is unfamiliar to her. Finally, she successfully associates the sound form [ˈɛkskəˌveɪtər] with the construction vehicle she has seen, though she has been only briefly exposed to the word-to-object association.

Despite the challenges involved in learning words, young children nevertheless become efficient word-learners within the first two years of life. They show recognition of some common words from six to nine months (Bergelson & Swingley, 2012), they usually utter their first words around the first birthday (Bloom, 2002), and their vocabulary size significantly increases during the second year of life especially from around 18 months to 24 months old, known as the "vocabulary spurt" period (Goldfield & Reznick, 1990).

Children learn words from input. It goes without saying that the novel word labels children hear from the input are crucial for children to learn words. Aside from *what* the word labels are, *how* these novel words are uttered may also matter for word learning. In the word-learning scene described above, the mother uses an exaggerated tone when she talks to her daughter. This speaking style is known as “infant-directed speech” (IDS), which caregivers typically use when addressing young children. Prototypical IDS has distinctive features compared to adult-directed speech (ADS). For example, IDS has more diminutives, hyperarticulated vowels, shorter utterances, more repetitions, and is grammatically less complex compared to ADS (see reviews in Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015; Soderstrom, 2007). The most salient feature of IDS is its exaggerated prosody. Such exaggerated prosody includes a higher pitch, a larger pitch range, and a slower speaking rate compared to ADS (see a review in Soderstrom, 2007). So far, extensive research has shown that IDS prosody differs from ADS across languages (e.g., Fernald et al., 1989; Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002), and there is some evidence to suggest that prototypical IDS facilitates American-English-learning children’s word learning as compared to ADS (Graf Estes & Hurley, 2013; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011). However, the role of prosodic input in early word learning has not yet been fully understood. Specifically, no study to date has addressed (1) the prosody of IDS in word-learning contexts in which mothers introduce unfamiliar words to children; (2) whether prototypical IDS facilitates word-to-object mapping across languages; and (3) whether there is a link between the IDS prosody specific to word-learning contexts and children’s language outcomes (e.g., vocabulary size and online word learning performance). Moreover, as most studies on IDS focus on a single language (mainly American English) (Frank et al., 2017), it is not clear whether the findings on English IDS can be generalized to typologically distant languages that differ in their prosodic characteristics. Crucially, cross-linguistic studies on the nature of IDS prosody specific to word-learning contexts as well as the potential facilitative effects of IDS on word learning are still lacking in the literature.

The overarching goal of this dissertation is to examine the role of prosodic input in children’s word learning. To achieve this goal, we conducted a cross-linguistic investigation of IDS in two languages that are typologically distinct: Dutch and Mandarin Chinese. The dissertation consists of five articles (Chapter 2 to Chapter 6). This introductory chapter (Chapter 1) provides a review of the literature, addresses methodological choices, introduces the main experiments, gives an overview of the five empirical articles, and outlines the dissertation.

1.1 Infant-directed speech

Researchers have used many terms to describe the linguistic environment in children's early language acquisition, such as "baby talk," "caregiver speech," "infant-directed speech," "child-directed speech," or more broadly, "linguistic input." Some of these terms are often used interchangeably and with some confusion (see Saxton, 2008 for a review). In the current dissertation, we mainly use the term "infant-directed speech" (IDS). The term "IDS" has two interpretations in the literature, which are illustrated as follows:

First, IDS is commonly used to refer to the prototypical speaking style caregivers use when addressing their child, and is mainly characterized by exaggerated prosody compared to adult-directed speech (ADS), i.e., speech addressed to adults. IDS is often studied in comparison with ADS to highlight the differences between the two speech registers. In this sense, IDS is also called "motherese," "parentese," or "baby talk." Not only mothers but also fathers, grandmothers, and other caregivers use this speaking style when talking to children (Fernald et al., 1989; Shute & Wheldall, 1999; Shute & Wheldall, 2001). Despite its name, the addressees of this speaking style are not limited to young infants. For example, speakers tend to modify their prosody in similar ways when talking to toddlers, pets, as well as foreigners (Knoll & Costall, 2015; Xu, Burnham, Kitamura, & Vollmer-Conna, 2013). Some researchers also make a distinction between speech addressed to infants (in a narrow sense, children younger than 12 months) and child-directed speech (CDS) (speech directed to older children). Indeed, the prosodic characteristics and functions of speech addressed to infants may differ from CDS (Liu, Tsao, & Kuhl, 2009). Nevertheless, we adopt a broad definition of infants (i.e., 0 to 2 years old) in this dissertation and the term "IDS" refers to speech addressed to children under the age of two.

Second, it should be noted that not all speech that is directed to children has an (equally) exaggerated prosody. There is considerable individual variation as well as variation dependent on the various daily activities (Ramírez-Esparza, García-Sierra, & Kuhl, 2017; Tamis-Lemonda, Custode, Kuchirko, Escobar, & Lo, 2018). As such, in some studies IDS is also used to refer to *any* type of speech that is directed to an individual child (e.g., Kalashnikova & Burnham, 2018; Weisleder & Fernald, 2013). This usage of the term is similar to the terms "caregiver speech" or "maternal input."

In the present dissertation, the term "IDS" can have either of these two interpretations depending on the context; we use "prototypical IDS" when highlighting its prosodic differences from ADS to avoid ambiguity.

1.2 Global IDS prosody

Many researchers have compared the prosody between ADS and prototypical IDS (see reviews in Cristia, 2013; Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015; Soderstrom, 2007). The approach taken in these studies has been to measure prosody—such as pitch properties and speaking rate—as a global property of IDS compared to ADS. The IDS samples are usually collected from mothers interacting with their children naturally at home or (semi-)structured mother-child interactions in laboratory settings. The ADS samples are typically conversations between a mother and an experimenter. Most studies have measured pitch properties such as mean pitch and pitch range, while relatively fewer studies have measured the temporal properties, including speech rate and articulation rate.

Summarizing these studies, compared to ADS, IDS in almost all languages exaggerates pitch. The measurement of pitch in these studies is fundamental frequency (F0), the main acoustic correlate of pitch. For instance, Fernald et al. (1989) compared IDS addressed to 10- to 12-month-old children and ADS in American English, British English, French, German, Italian, and Japanese. Their results showed that mothers had a higher mean pitch, a higher minimum pitch and maximum pitch, larger pitch range, and greater pitch variability (the standard deviations of the mean F0) when addressing children compared to adults in all languages under investigation. Similar results have been obtained in a range of languages including Australian English, Cantonese, Korean, Mandarin Chinese, Sri Lankan Tamil, Tagalog, and Thai, among others (Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002; Liu et al., 2009; Narayan & McDermott, 2016; Xu Rattanasone, Burnham, & Reilly, 2013). One observed exception is Quiché Mayan. In this language there is little evidence of pitch modifications in IDS (Ingram, 1995; Bernstein Ratner & Pye, 1984; though see Pye, 1986 for non-prosodic features of Quiché IDS including whispering, IDS verb forms, etc.).

Even though it is widely accepted that IDS is slower than ADS, only a few studies have specifically measured the speaking rate of IDS. Some studies show that IDS is slower than ADS in various languages. For example, German IDS addressed to newborns had a slower articulation rate (excluding pauses, measured in syllables per second) than ADS (Fernald & Simon, 1984). American English addressed to toddlers (17 to 20 months of age) had a slower speech rate (including pauses) in comparison with ADS (Bernstein Ratner, 1985). Dutch IDS addressed to 11-month-old children had a slower speech rate than ADS (Johnson, Lahey, Ernestus, & Cutler, 2013; Van de Weijer, 1999). Similar results were found in Cantonese, Korean, Swedish, and Tagalog: In all these languages, IDS has been reported to be slower than

ADS (Narayan & McDermott, 2016; Sjons, Hörberg, Östling, & Bjerva, 2017; Tang & Maidment, 1996). However, opposing evidence exists. First, IDS does not seem to be slower than ADS in some languages such as Sri Lankan Tamil (Narayan & McDermott, 2016). Second, even for languages in which IDS is globally slower than ADS, IDS does not seem to be slower than ADS across entire utterances. For example, Canadian English IDS was not slower than ADS when excluding utterance-final syllables (Church, Bernhardt, Shi, & Pichora-Fuller, 2005). Also, Japanese IDS only slowed down at phrase-final and utterance-final positions (Martin, Igarashi, Jincho, & Mazuka, 2016). It should be noted that the methods for measuring speaking rate of IDS varied across the studies. Some of these studies measured *speech rate*, which includes silent pauses, while others measured *articulation rate*, excluding silent pauses. Considering the cross-linguistic differences as well as the diverse measurements used in these studies, it is not conclusive whether slow speaking rate is a common characteristic of IDS across languages.

Taken together, there is robust evidence of exaggerated pitch in IDS as compared to ADS at the global level. In particular, IDS has a higher mean pitch and a larger pitch range as compared to ADS in a diversity of languages. The empirical findings for the temporal aspect of IDS are rather mixed. There is no conclusive evidence on whether IDS is slower than ADS across languages or across entire utterances.

1.3 IDS in prosodically distinct languages

1.3.1 Language universality and language specificity in IDS

As reviewed above, the exaggerated prosody of IDS is found in almost all languages with only one observed exception, namely Quiché Mayan (Ingram, 1995; Bernstein Ratner & Pye, 1984). As a result, the exaggerated prosody of IDS is often considered as a universal feature across languages and cultures (Bryant & Barrett, 2007; Fernald et al., 1989; Grieser & Kuhl, 1988).

However, language specificity of IDS prosody exists, namely in two aspects. First, the degree of pitch modification varies among languages. That is, not all IDS is exaggerated to the same extent. For instance, the difference in mean pitch between American English ADS and IDS is larger than that in British English, French, German, Italian, or Japanese (Fernald et al., 1989). In the same vein, Australian English IDS is more exaggerated in mean pitch and pitch range than Thai IDS compared to ADS (Kitamura et al., 2002).

Second, it is not clear whether IDS speaking rate is consistently slower than ADS across languages and across utterances. As previously reviewed, even though IDS is slower than ADS in languages such as English and German, there is no evidence of slowing down in Sri Lankan Tamil IDS (Narayan & McDermott, 2016). In addition, several studies suggest that IDS is not slower than ADS across entire utterances in Canadian English and Japanese (Church et al., 2005; Martin et al., 2016).

Even though there has been some evidence of language universality and language specificity in IDS within the literature, most studies on IDS have focused on a single language instead of taking more than one language into consideration. Also, few studies have addressed how language specificity of IDS, if it exists, might be related to prosodic differences among languages.

Recent research proposes that the prosodic modifications of IDS may be influenced by prosodic typology (see Wang, Seidl, & Cristia, 2016 for a review). Languages differ in their prosody on various dimensions. For example, tonal languages such as Mandarin Chinese and Thai use pitch (e.g., pitch height and pitch contour) to distinguish lexical meanings. In contrast, pitch at the lexical level does not change word meanings in non-tonal languages (e.g., English and Dutch). Languages such as Dutch and English have lexical stress instead. In stress languages, lexical stress is placed on a given syllable in a word.

In addition to the differences in word prosody, languages differ in their rhythmic class. Traditionally, three rhythmic classes have been proposed: stress-timed, syllable-timed, and mora-timed (Abercrombie, 1967; Beckman, 1992; Grabe & Low, 2002; Ladefoged, 1975). It was proposed that in stress-timed languages, stressed and unstressed syllables are distinguished in terms of duration and syllable weight, while syllable-timed languages have nearly equal weight and time in all syllables. Also, the durational variability is greater in stress-timed languages compared to syllable-timed languages (Grabe & Low, 2002). Even though this classification has remained controversial, and durational differences between these languages are gradient rather than categorical, acoustic evidence suggests that English and Dutch are closer to the typological extreme of stress-timed languages (e.g., Ramus, Dupoux, & Mehler, 2003), while Mandarin Chinese is closer to the typological extreme of syllable-timed languages (e.g., Mok & Dellwo, 2008).

As previously reviewed, IDS is characterized by generally exaggerated prosody, including pitch and temporal modifications. Thus, it is possible that tone (mainly marked by pitch) and stress (mainly realized by duration) at the word level, as well as speech rhythm at the utterance level, interact with global pitch or temporal modifications of IDS. Such interaction may have two consequences. First, such interaction may constrain the generally exaggerated prosody in IDS, potentially contributing to cross-linguistic variation in global IDS prosody. For example, as Kitamura et al., (2002, p. 386) pointed out, their finding that mean pitch was lower and pitch range was smaller in both Thai ADS and IDS compared to Australian

English may in part be attributed to lexical tone, as Thai mothers restricted pitch changes so that lexical tones in IDS were identifiable. Second, prosody at the word level, for example, lexical tones, may be *hyperarticulated* (exaggerated) or *hypoarticulated* (distorted) due to the global prosodic modifications (see Section 1.3.3 for more discussion). Whether these possibilities hold, and how these interactions may take place, requires cross-linguistic investigations on languages that are typologically distinct in word prosody. Specifically, a cross-linguistic investigation of tone vs. non-tone languages or stress vs. non-stress languages may contribute to a better understanding of this issue. To our knowledge, so far only one study has directly compared IDS in a tonal language (Thai) and a non-tonal language (Australian English) (Kitamura et al., 2002). As such, more cross-linguistic investigations of pairs of typologically distinct languages are needed to better determine the language-universal and language-specific aspects of IDS.

1.3.2 Previous studies on Dutch and Mandarin Chinese IDS

In the current dissertation, we selected Dutch and Mandarin Chinese—two languages that are typologically distinct and differ in their prosodic characteristics. First, Dutch and Mandarin Chinese have different syllable complexity. For example, Dutch allows a complex syllable structure: (C)(C)(C)V(V)(C)(C)(C)(C) (Van der Hulst, 1984), whereas Mandarin Chinese has a relatively simpler syllable structure (C)(G)V(X) in which G is a glide and X can either be a consonant or vowel (Bao, 1996). Second, Dutch and Mandarin Chinese have different word prosody: Dutch has lexical stress while Mandarin Chinese has lexical tones. In Dutch, the main acoustic correlates of lexical stress are duration and spectral balance (Sluijter & Van Heuven, 1996). In terms of speech rhythm, Dutch is typically considered a stress-timed language (Grabe & Low, 2002; Ramus, Dupoux, & Mehler, 2003) and Mandarin Chinese is a syllable-timed language (Mok & Dellwo, 2008). In both languages, prosody can be used to mark contextually new information (or focus). Dutch speakers usually produce focal words with a larger pitch range and a longer duration compared to non-focal words (e.g., Chen, 2009; Hanssen, Peters, & Gussenhoven, 2008). In Mandarin Chinese, contextually new information has a larger pitch range, a longer duration, and a higher mean pitch compared to contextually given information (e.g., Chen & Braun, 2006; Chen & Gussenhoven, 2008; Xu, 1999).

Even though there are many existing studies on English IDS, especially American English IDS, to this day only a few empirical studies have investigated Dutch and/or Mandarin Chinese IDS. With respect to Dutch, Van de Weijer (1999) examined IDS addressed to a Dutch-German bilingual child from 6 to 9 months of age in a case study. The mother of the child was a German native speaker who spoke both Dutch and German to her child. These two languages were not separated in the analysis. The results showed that utterances in mothers' IDS had a higher mean F0, a higher minimum F0, a higher maximum F0, a larger F0 range, and a slower speech rate (including utterance-internal silent pauses) as compared to her ADS, in accordance with the universal prosodic exaggeration typical of IDS. Benders (2013) further investigated Dutch IDS addressed to 11- and 15-month-old children ($N = 18$; longitudinal design). The pitch measures were median F0 and F0 excursions (F0 range divided by utterance duration) at the utterance level. The results showed that median F0 was higher and F0 excursions were larger in IDS compared to ADS at both ages. In addition to examining the prosody of Dutch IDS, Benders (2013) also found that vowels were *hypoarticulated* in Dutch IDS, inconsistent with results from languages such as American English, which has *hyperarticulated* vowels in IDS (e.g., Kuhl et al., 1997). Johnson et al. (2013) described an audio and video corpus of Dutch ADS and IDS addressed to 28 11-month-old children. The caregivers included one parent and an additional caregiver (the other parent or grandmother). In this study, IDS was collected during a free caregiver-child interaction session and a word-teaching task. In the word-teaching task, the caregiver was instructed to teach four words (including a proper name, a common noun, an adjective, and a verb) to their children. Preliminary analyses showed that utterances in IDS were shorter and slower compared to ADS. In the word-teaching task, caregivers produced adjectives less frequently compared to common nouns, proper nouns, or verbs. Pitch properties were not included in their analyses.

Grieser and Kuhl (1988) provided the first empirical evidence of prosodic modifications of Mandarin Chinese IDS. In this study, IDS addressed to infants aged 6 weeks to 10 weeks was examined. The results showed that, consistent with previous findings on non-tonal languages such as English and German, Mandarin Chinese IDS was also characterized by a higher mean pitch and a larger pitch range compared to ADS. Aside from this study which measured the prosodic modifications of IDS, several studies examined the segmental (e.g., vowels) and suprasegmental (i.e., lexical tones) properties in Mandarin Chinese IDS. For example, Liu et al. (2009) investigated the acoustic features of mothers' speech addressed to preverbal children (7 to 12 months) in Taiwanese Mandarin (a variation of Mandarin Chinese). Instead of measuring pitch at the utterance level, this study measured the mean F0 and F0 range on the vowels and their results showed that IDS had a higher mean F0 and F0 range than ADS. Tang, Xu Rattanasone, Yuen, and Demuth (2017) examined the vowels and lexical tones in Mandarin Chinese IDS addressed to 12-month-old children, showing that while vowels were hyperarticulated in IDS, lexical tonal hyperarticulation was only observed in utterance-final position in IDS.

Based on the limited number of studies on Dutch and Mandarin Chinese IDS, we can conclude that both Dutch and Mandarin Chinese IDS have exaggerated prosody, characterized by an overall higher mean pitch and a larger pitch range compared to their ADS counterparts. The global pitch modifications of IDS are consistent with the majority of languages, but no study has examined the speaking rate of Mandarin Chinese IDS. In addition, among these studies, the speech materials varied and the age groups under investigation differed. Furthermore, to date there exists no cross-linguistic study including both Dutch and Mandarin Chinese.

1.3.3 Lexical tones in IDS in tonal languages

As previously mentioned, Mandarin Chinese is a tonal language in which pitch is employed to differentiate lexical meanings (Yip, 2002). Mandarin Chinese has four lexical tones: Tone 1 (high-level), Tone 2 (mid-rising), Tone 3 (low-dipping), and Tone 4 (high-falling) (Chao, 1948; Howie, 1976). The meaning of a word can change depending on its tone. For example, *mā* (Tone 1) means "mother", *má* (Tone 2) means "hemp", mǎ (Tone 3) means "horse", and mà (Tone 4) means "to scold". As such, tonal information is essential to early word learning for tone language learners: a child learning a tone language such as Mandarin Chinese must be able to distinguish lexical tones when learning words.

In IDS, lexical tones interact with a globally exaggerated prosody at the intonational level. As a result, lexical tones may either be *hyperarticulated* (or exaggerated) or *hypoarticulated* (or distorted) due to such interaction. So far only a few studies on various languages have investigated the articulation of lexical tones in IDS, and the results are inconsistent among languages. Some studies suggest that lexical tones in IDS are *hypoarticulated* (Kitamura et al., 2002 on Thai; Papoušek & Hwang, 1991 on Mandarin Chinese) which may hinder children's acquisition of tones, while others suggest that lexical tones in IDS are in fact *hyperarticulated* (e.g., Liu, Tsao, & Kuhl, 2007 on Taiwanese Mandarin; Tang, Xu Rattanasone, Yuen, & Demuth, 2017 on Mandarin Chinese), which may potentially support children's discrimination of tonal contrasts and further facilitate word learning.

It should be noted that these studies measured lexical tones in IDS in the first year of life, before children's vocabulary starts to increase rapidly at around 16 to 18 months (Goldfield & Reznick, 1990). Studies have shown that the degree of lexical tonal hyperarticulation in both Thai and Cantonese decreases as children reach 12 months (Kitamura et al., 2002; Xu Rattanasone, Burnham, & Reilly, 2013). As lexical tone is crucial to word meanings in Mandarin Chinese, it remains a question whether lexical tones are hyperarticulated during the "vocabulary spurt" period from 18 to 24 months in order to facilitate word learning. Furthermore, whether there are age-related changes of tone hyperarticulation in the second year is also unclear.

Thus, in addition to addressing the word and utterance prosody of IDS in word learning, we investigated whether lexical tones in Mandarin Chinese IDS are hyperarticulated during the second year of life, and whether lexical tones change from 18 months to 24 months.

1.4 Potential linguistic functions of IDS

Thus far, we have reviewed literature regarding the distinctive prosody of IDS. The question remains: What are the functions of IDS prosody? Researchers have proposed three functions of IDS.

First, compared to ADS, IDS attracts infants' attention. For example, Cooper and Aslin (1990) found that English-learning newborns prefer listening to IDS stimuli which are characterized by a higher mean F0, a larger F0 range, a slower articulation rate, and longer pauses, compared to ADS stimuli. Similar results were obtained from English 4-month-old children (Fernald, 1985) and Cantonese 4.5- and 9-month-old children (Werker, Pegg, & McLeod, 1994). Segal and Newman (2015) found that 16-month-old English-learning infants still prefer prototypical IDS prosody as compared to ADS. Moreover, children's listening preference for IDS may change with age, Hayashi, Tamekawa, and Kiritani (2001) found that while Japanese infants' preference for IDS was strong at 4 to 6 months, there was no evidence to support a preference for IDS at 7 to 9 months, while listening preference for IDS was shown again at 10 to 14 months. Taken together, there is robust evidence to suggest that infants prefer listening to prototypical IDS over ADS, which has been confirmed by a meta-analysis (Dunst et al., 2012).

Second, prototypical IDS conveys positive affect. Trainor, Austin, and Desjardins (2000) suggested that the exaggerated prosody of IDS is similar to that of ADS specifically with a positive affect. Furthermore, children's listening preference for IDS is mainly attributed to the positive affect of IDS. In a study that examined 6-month-old English-learning children's listening preference, Singh, Morgan, and Best (2002) found that the exaggerated prosody of IDS did not elicit infants' preference; in fact, positive affect played an important role in attracting children's listening preference.

Third, prototypical IDS has been claimed to facilitate language acquisition as compared to ADS beyond attracting attention and conveying positive affect. To be sure, the attentional and emotional effects of IDS may *indirectly* contribute to language acquisition, however, whether IDS has specific functions that *directly* support language acquisition remains controversial.

First, it is unclear whether certain aspects of IDS are organized in a way that may directly serve linguistic purposes. In order to answer this question, studies have focused on whether segmental categories such as vowels are hyperarticulated across languages, as vowel hyperarticulation may potentially support infants' speech perception and category learning (e.g., Adriaans & Swingley, 2017). In an oft-cited study, Kuhl et al. (1997) found that vowels in American English, Russian, and Swedish IDS addressed to 2- to 5-month-old children were consistently hyperarticulated, indicating that vowels were more distinct from each other in the IDS of these languages compared to in ADS. Similarly, vowels were shown to be hyperarticulated in French (Dodane & Al-Tamimi, 2007) and Mandarin Chinese IDS (Liu, Kuhl, & Tsao, 2003; Tang et al., 2017). While there has been some evidence supporting vowel hyperarticulation in IDS, conflicting results are emerging. Vowel hyperarticulation does not seem to exist in languages such as Cantonese (Xu Rattanasone, Burnham, & Reilly, 2013). Moreover, vowel *hypoarticulation* was shown in Dutch (Benders, 2013), Japanese (Miyazawa, Shinya, Martin, Kikuchi, & Mazuka, 2017) and Norwegian (Englund, 2017). Overall, whether segmental properties of IDS are hyperarticulated in a way that may facilitate children's speech perception across all languages requires further investigation. In addition to vowel hyperarticulation, a few studies have examined the prosodic marking of contextually new information in IDS, which may support children's word learning (to be reviewed in detail in Section 1.5 of this dissertation).

Second, a range of studies have shown that prototypical IDS facilitates English-learning children's online word processing, including word segmentation, word recognition, and word-to-object mapping, as compared to ADS (e.g., Graf Estes & Hurley, 2013; Ma et al., 2011; Mani & Pätzold, 2016; Singh, Nestor, Parikh, & Yull, 2009; Thiessen, Hill, & Saffran, 2005). However, the evidence has been largely based on American English without extending to other languages. In these studies, children hear auditory stimuli that have similar speech content but are produced with either ADS or prototypical IDS prosody: IDS stimuli have a higher pitch, larger pitch range, and a slower speaking rate compared to ADS stimuli. Children's online word processing performances are compared between ADS and IDS conditions. Specifically, IDS has been shown to support seven-month-old English- and German-learning infants' word segmentation compared to ADS (Thiessen, Hill, & Saffran, 2005; Mani & Pätzold, 2016). Children could only segment words in continuous speech from IDS input but not when hearing ADS input. Also, English-learning infants could only successfully recognize the words that they were familiarized with in prototypical IDS, but not if the words were initially familiarized in ADS (Singh, Nestor, Parikh, & Yull, 2009).

Word learning, in a broad sense, requires the previously mentioned tasks such as word segmentation, word recognition, as well as word-to-object mapping. In the current dissertation, we focus on the effect of IDS in a narrow sense of word learning: word-to-object mapping. In studies on word-to-object mapping, children need to associate a novel word label with a novel object. Previous studies suggest that American English IDS facilitates children's word-to-object mapping at least before 21 months of age. In particular, 17.5-month-old and 21-month-old children could only reliably learn word-to-object mapping reliably when the words were produced in prototypical IDS but not in ADS, and only after 27 months could children reliably learn word-to-object mapping in the ADS condition (Graf Estes & Hurley, 2013; Ma et al., 2011).

The findings from Graf Estes and Hurley (2013) and Ma et al. (2011) together suggest that prototypical IDS facilitates American English children's word-to-object mapping as compared to ADS. However, as American English IDS seems to be the most exaggerated IDS among languages (Fernald et al., 1989), it remains unclear whether these findings on American English IDS can be generalized to other languages or even another variation of English. Previous studies on word segmentation have shown that while American-English-learning children could segment words from continuous speech as early as 7.5 months (Jusczyk & Aslin, 1995), British-English-learning children failed to do so in the same paradigm until 10.5 months old. Interestingly, 10.5-month-old British English children could only succeed in word segmentation tasks when the stimuli was particularly exaggerated in prosody (Floccia et al., 2016). As the authors suggested, one possible explanation for the inconsistent results between American English and British English is the difference in the degree of prosodic exaggeration in IDS between these two English dialects. So far there has been no evidence of the facilitative effects of prototypical IDS on word-to-object mapping in languages other than American English, it is certainly possible that previous findings on word-to-object mapping in American English children may not necessarily apply to other languages.

In addition, there is some evidence to suggest that the distinctive prosody of IDS (alone) may not account for the facilitative effects. In the studies reviewed above, the only difference between ADS and IDS stimuli was global prosody. However, Song, Demuth, and Morgan (2010) took a further step by examining which acoustic properties of IDS facilitated children's word recognition. Their results indicate that vowel hyperarticulation and slow speaking rate, but not pitch range, facilitate 19-month-old American English children's word recognition. Also, in a meta-analysis, Spinelli, Fasolo, and Mesman (2017) suggest that the evidence from existing studies supports the view that IDS prosody has more attentional functions and less linguistic functions especially for infants before the age of one.

Aside from directly comparing children's word learning performances in ADS and prototypical IDS at the group level, some studies have explored the correlations between individual IDS prosody and children's language outcomes. Even though the correlations do not necessarily entail any facilitative effects of IDS, these studies have provided insight into the links between individual IDS and children's language outcomes (see Section 1.7 for more discussion).

Taken together, whether IDS has linguistic functions beyond attentional and emotional functions, and whether it is the prototypical prosody of IDS that accounts for the potential linguistic functions, remain unclear.

1.5 IDS prosody specific to word-learning contexts

As illustrated above, a large amount of literature has compared the global prosody of IDS with that of ADS, with some evidence to suggest that prototypical IDS may facilitate children's online word-to-object mapping. However, studies comparing the prosody of these two speech registers usually view the exaggerated prosody of IDS as a global property without zooming in on a specific context such as word-learning contexts in which mothers introduce unfamiliar words to children. Such word-learning contexts arguably provide the most direct input for children to learn words. The question remains: is the prosody of IDS specific to word-learning contexts modified in a way that may potentially support word learning? Based on Hyper and Hypo-speech (H&H) theory (Lindblom, 1990), Fernald (2000, p. 250) proposed that "caretakers also arrange the words they speak to infants in ways which may facilitate comprehension." It is possible that this proposal can be further extended to word learning. If mothers adapt their speech prosody according to their children's vocabulary knowledge, they may make a distinction when they introduce familiar vs. unfamiliar words to children by modifying their prosody. However, no research to date has specifically examined IDS prosody specific to word-learning contexts in which mothers introduce unfamiliar words to children as compared to familiar words. Instead, the most relevant studies in the literature focus on the prosodic marking of contextually new information as compared to contextually given information in English IDS (Bortfeld & Morgan, 2010; Fernald & Mazzie, 1991; Fisher & Tokura, 1995; to review later).

It is important to distinguish these two contrasts: (1) familiar words vs. unfamiliar words; and (2) contextually new information vs. contextually given information. Young children, as early language learners, encounter unfamiliar words on a regular basis, whereas words directed to adults are usually only "contextually new" within a specific conversational context and rarely novel or unfamiliar to the addressee. In the example illustrated at the beginning of this chapter, the mother introduces the word "excavator" to her child and repeats the word form "excavator" several times, as the girl does not yet know the word. Thus, the word "excavator" is unfamiliar to the child. However, this conversation would rarely happen between a mother and another adult. Usually she would presume that the other adult addressee already knows the word "excavator." When she talks about the word "excavator" for the first time with another adult, this word may be "contextually new" in discourse but is not necessarily "unfamiliar." Thus, in this dissertation the term "unfamiliar words" is used to refer to words that children have not yet acquired, and "contextually new information" refers to parts of an utterance that are being introduced to the conversation for the first time but are not necessarily unfamiliar to the addressee.

In ADS, speakers usually highlight contextually new information in discourse by increasing pitch and/or enlarging pitch range, while downplaying given information by reducing these prosodic parameters. Additionally, they use pronouns in the place of lexical forms when mentioning a referent for the second time (Chafe, 1976; Gundel, 1999; Halliday, 1967; Krifka & Musan, 2012). In IDS, however, mothers usually repeat the same word several times when talking to children instead of replacing the word with a pronoun (Fernald & Morikawa, 1993; Fernald & Simon, 1984).

As previously mentioned, existing studies have only examined the prosodic marking of contextually new information when (English-speaking) mothers address children vs. adults. Fernald and Mazzie (1991) found that when American-English-speaking mothers addressed 14-month-old children, they tended to place prosodic prominence on both contextually new and contextually given words in IDS. Specifically, mothers had a greater tendency of placing the F0 peak of the utterance on both contextually new words and contextually given words in IDS as compared to ADS. Also, mothers increased the maximum F0 on contextually given words, a speech pattern that did not hold in ADS. Fisher and Tokura (1995) also examined the prosody of contextually new and contextually given words when mothers address their child. Their results showed that the first mention (contextually new) received prosodic prominence as compared to the second mention (contextually given), shown by a higher pitch, a larger pitch range, and a longer duration of the first mention. Such given-new contrast was consistent with ADS. Bortfeld and Morgan (2010) extended

Fisher and Tokura (1995) to multiple mentions, and their results showed a significant quartic trend in four prosodic measures: mean F0, maximum F0, F0 range, and duration, suggesting that mothers alternated between stressed and unstressed realizations across multiple mentions in American English IDS.

The studies reviewed above were only concerned with the prosodic differences between contextually new and contextually given words in IDS. Even though results from these studies were all interpretable as evidence for the highlighting of "new" words in IDS, which may have implications for word learning, none of these studies have specifically taken the familiarity of words into account and examined whether mothers use prosody to differentiate unfamiliar words and familiar words in IDS. Previous studies suggest that mothers have a good knowledge of children's vocabulary at an item level (Fenson et al., 2007; Styles & Plunkett, 2009). In other words, mothers are aware of what words children know and what children have not yet understood. However, no study has examined whether mothers integrate this knowledge in speech to children. In order to better understand the prosodic input in word-learning contexts, it is important to take familiarity of words into consideration and investigate whether mothers use prosody to distinguish unfamiliar words from familiar words in IDS. Even if the function of global prosody of IDS may mainly be attentional, investigations into the prosody of IDS in word-learning contexts may reveal whether IDS prosody serves linguistic purposes. Moreover, as languages differ in their prosody, IDS in different languages may use different prosodic cues in word-learning contexts depending on their word prosody or rhythmic properties. As such, in addition to investigating whether mothers use prosody to distinguish unfamiliar words from familiar words, this study included Dutch and Mandarin Chinese in order to understand whether the role of familiarity on prosody in word-learning contexts manifest itself differently in prosodically distinct languages.

1.6 Age effect

In addition to potential cross-linguistic differences and the role of familiarity, a child's age may further influence the prosody of IDS and the prosody of IDS specific to word learning. A few studies have explored the age-related changes of the pitch and temporal cues of IDS. A general trend of the age-related changes is that IDS becomes more similar to ADS as children grow older and gain better linguistic competence (e.g., Kitamura et al., 2002; Xu Rattanasone, Burnham, & Reilly, 2013), though such a trend may not be linear and may not hold in all cases.

Regarding pitch, Stern, Spieker, Barnett, and MacKain (1983) showed that pitch properties in IDS were more exaggerated when mothers addressed 4-month-old children as compared to older children. Kitamura et al. (2002) found that in Australian

English, the mean F0 of IDS increased at 6 months, decreased at 9 months, and then increased again at 12 months. The degree of pitch exaggeration in Taiwanese Mandarin was larger when addressing preverbal children compared to five-year-old children (Liu et al., 2009). However, no evidence of age-related changes in pitch was found in Korean, Tagalog, or Sri Lankan Tamil IDS addressed to 4- to 16-month-olds (Narayan & McDermott, 2016).

Only two studies have examined whether the speaking rate of IDS changes with age. Even though Narayan and McDermott (2016) found no evidence of age-related changes in pitch, they showed that the speaking rate of IDS in Korean ($N = 6$), Tagalog ($N = 5$), and Sri Lankan Tamil ($N = 5$) overall increased from 4 to 16 months, and was similar to ADS at 16 months. On the contrary, a longitudinal study on Swedish IDS revealed that IDS became faster from 7 to 33 months, but it was still slower than ADS when children were 33 months old (Sjons et al., 2017).

In the first year of life, children's speech perception has tuned into their first language(s), known as "perceptual reorganization" or "perceptual attunement" (Werker & Tees, 1984; Werker, 2018). In the literature reviewed above, no study has focused on the age-related changes of IDS in the second year of life, when children have already undergone "perceptual reorganization" and have started to become efficient word learners. Children's lexical development starts as early as six to nine months, at which time they can already recognize some common words such as food and body parts (Bergelson & Swingley, 2012). At the end of the first year, children usually utter their first words (Bloom, 2001). Later, children's vocabulary size significantly increases from around 16 to 18 months onwards, known as the "vocabulary spurt" (Goldfield & Reznick, 1990). In addition to the increase in vocabulary size, children's "fast mapping" ability—the ability to map a novel label to a novel object based on minimal exposure—also gradually improves during the same period. In particular, 18-month-old children do not reliably prefer mapping a novel label to a novel object as compared to a familiar object, but 24-month-old children can reliably associate a novel label to a novel object rather than a familiar object (Bion, Borovsky, & Fernald, 2013).

The current study, therefore, specifically targets 18 months and 24 months, a crucial period for word learning, and asks whether the global prosody of IDS and the prosody of IDS specific to word learning change during this period.

1.7 The link between prosodic input and language outcomes

Prototypical IDS has been shown to facilitate word-to-object mapping, at least for American-English learning children under 21 months of age (Graf Estes & Hurley, 2013; Ma et al., 2011). However, simply comparing children's word learning performance in response to prototypical ADS and IDS at the group level failed to address the relationship between the prosody of individual IDS and children's language outcomes (e.g., vocabulary size and online word learning performance).

It is well established that the *quantity* of IDS is correlated with children's vocabulary size. Children who hear more words in mother-child interactions tend to have larger vocabularies (Hart & Risley, 1995; Hoff, 2006; Hoff & Naigles, 2002; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Mahr & Edwards, 2018). Also, many aspects of input *quality* are associated with children's vocabulary size. For example, lexical richness, syntactic complexity, repetitiveness (indicated by type-token ratio), the number of "rare words" in the input, and vowel hyperarticulation have all been shown to correlate with children's vocabulary size (Hoff & Naigles, 2002; Hartman, Bernstein Ratner, & Newman, 2016; Liu et al., 2003; Newman, Rowe, & Bernstein Ratner, 2015; Rowe, 2012).

Surprisingly, even though prosodic modifications, including pitch modifications and temporal modifications, are among the most distinctive features of IDS as compared to ADS, the relationship between individual IDS prosody and children's vocabulary size is often ignored in previous studies. In fact, only a few studies have investigated whether the prosody of IDS is correlated with children's vocabulary, and the results are conflicting. For example, Porritt, Zinser, Bachorowski, and Kaplan (2014) found that American-English-learning infants' expressive vocabulary (percentile scores) was positively correlated with F0 range in their mothers' IDS. However, in a recent study, Kalashnikova and Burnham (2018) found no evidence that the extent of pitch exaggeration in Australian English IDS is correlated with children's vocabulary size. Similarly, Song, Demuth, and Morgan (2018) did not find any correlations between the prosody of individual mothers' IDS (mean pitch and pitch range) at 17 months and children's vocabulary size at 19 or 25 months. As for the temporal cues, one study has shown that a slow speaking rate in IDS at seven months significantly predicts larger expressive vocabulary at two years of age (Raneri, 2015).

The predictors in these studies mainly involve the globally exaggerated prosody of IDS. However, it is possible that instead of this globally exaggerated prosody, it is the particular ways in which mothers introduce unfamiliar words that are associated with children's vocabulary size. In the current dissertation, we ask whether the prosody of IDS specific to word-learning contexts is correlated with children's vocabulary size.

In correlation studies relating IDS properties and language outcomes, vocabulary size is commonly used as a measure of language outcomes. Another more direct measure of word-learning ability would be children's online word processing ability (as measured in an experimental setting). So far, only a few studies have shown positive correlations between the quantity of input and children's speed of online word recognition (e.g., Hurtado, Marchman, & Fernald, 2008). In addition to vocabulary size and online word recognition, yet a third possible language outcome that might depend on IDS is children's online ability for word-to-object mapping. However, no study has specifically examined whether there is a correlation between the prosodic quality of individual IDS prosody and children's online word-to-object mapping performance.

As previously mentioned, two studies to date have shown that prototypical IDS facilitates American English children's word-to-object mapping (Graf Estes & Hurley, 2013; Ma et al., 2011). However, these two studies have focused on the group differences without further exploring the role of input in the individual differences. Thus, it is not clear from these studies whether children's individual variation in their word-to-object mapping could be linked to their individual prosodic input. In this dissertation, we further explore whether individual children's online word-to-object mapping performances in ADS and IDS are correlated with their mothers' IDS prosody.

Two issues should be noted when interpreting the significant correlations between input quantity/quality and children's language outcomes found in the literature. First, these significant correlations have usually been interpreted as evidence for potential facilitative effects of IDS on children's language development. However, they could also be interpreted from a different perspective such that mothers might adapt their IDS based on their own observations of their children's linguistic knowledge. The alternative interpretation has often been ignored in previous studies. Second, even though both concurrent and longitudinal correlations have been explored in the studies discussed above, children's vocabulary size at an earlier age has rarely been taken into consideration in the analyses. As such, it is difficult to draw any conclusions as to whether IDS predicts children's vocabulary *growth* from these studies.

1.8 Research gaps, the main purpose of this dissertation, and experiments

Drawing from the literature reviewed above, five main research gaps are identified with regard to the role of prosodic input in word learning. First, no study to date has investigated the prosody of IDS specific to word-learning contexts. Second, studies on the prosody of IDS often focus on a single language without cross-linguistic investigations of prosodically distinct languages. Third, relatively few studies have examined IDS addressed to children during the "vocabulary spurt" period from 18 to

24 months. Fourth, there is no sufficient evidence that prototypical IDS facilitates children's word-to-object mapping across languages. Finally, it remains unclear whether there are correlations between the prosody of IDS and children's language outcomes, including vocabulary size and online word-to-object mapping performance.

The main purpose of this dissertation is to better understand the role of prosodic input in children's word learning. To achieve the main purpose and address the research gaps, we conducted three experiments including two production experiments on Dutch and Mandarin Chinese, respectively, as well as a word learning experiment on Dutch children. These experiments are reported in five empirical chapters (Chapters 2–6). Chapters 2, 3, and 4 focus on various aspects of the prosodic characteristics of IDS in word learning contexts. Chapters 5 and 6 target the links between prosodic input in word learning contexts and children's language outcomes.

In previous studies comparing the prosody of ADS and IDS, the speech samples of IDS are usually recorded during natural mother-child interactions at home or semi-structured mother-child interactions in laboratory settings, while the speech samples in the ADS condition are usually free conversations or interviews between the mother and an experimenter. As a result, the content and contexts usually differ between the two conditions as well as among studies. Thus, instead of comparing natural mother-child interactions with spontaneous adult speech, we elicited semi-spontaneous speech for both ADS and IDS. Specifically, in addition to simply matching the speech elicitation methods in the two speech registers we also used similar materials in the Dutch and Mandarin Chinese experiments. The experiments are summarized as follows:

In **Experiment 1**, Dutch-speaking dyads participated in a semi-spontaneous storybook-telling task when the children were 18 months and 24 months. To establish a word-learning context, the storybook contained words that were familiar or unfamiliar to the children. The Dutch experiment had a longitudinal design.

In **Experiment 2**, Mandarin-Chinese-speaking mothers of 18- and 24-month-old children participated in a semi-spontaneous storybook-telling task. The speech elicitation methods and materials were similar to Experiment 1. The major difference between Experiment 1 and Experiment 2 is that Experiment 2 had a cross-sectional design instead of a longitudinal design.

Experiment 3 is a word learning experiment adapted from Ma et al. (2011) using an Intermodal Preferential Looking Paradigm (Hirsh-Pasek & Golinkoff, 1996), in which children's word-to-object mapping was compared in ADS and IDS conditions. The major difference from Ma et al. (2011) was that, instead of a between-subject design, we adopted a within-subject design in the current study. The participants in this experiment were a subset of the participants in Experiment 1. These participants participated in Experiment 1 at both 18 months and 24 months, and were also tested on their word learning performances in ADS and IDS when they were 24 months of age.

In addition, we collected vocabulary information from all participants in the three experiments. All mothers filled in Dutch (N-CDI; Zink & Lejaegere, 2002) or Mandarin Chinese (M-CDI; Tardif, Fletcher, Liang, & Kaciroti, 2009) versions of MacArthur-Bates Communicative Development Inventories (CDI; Fenson, Bates, Dale, Marchman, & Reznick, 2007).

Experiment 1 examined the prosody of Dutch IDS in word-learning contexts. The results of Experiment 1 are reported in Chapters 2, 3, 5, and 6, which focus on different aspects of the speech data. Experiment 2 examined the prosody of Mandarin Chinese IDS in word-learning contexts. The results are reported in Chapters 2, 3, 4, and 5, focusing on different aspects of the speech data. As a result, there is some overlap in the introduction as well as the methods sections in these chapters. Experiment 3, a word learning experiment, is reported on in Chapter 6. The following section describes the research questions in each empirical chapter.

1.9 Chapter overview and research questions

Chapters 2 and 3 focus on the prosodic input in word-learning contexts. Specifically, Chapter 2 reports the temporal aspects of IDS, namely speaking rate, and Chapter 3 describes the pitch properties of IDS, including mean pitch and pitch range. Three main research questions are raised in these two chapters:

First, as no study to date has investigated the prosody of IDS specific to word-learning contexts in which mothers introduce unfamiliar words to children and the only relevant studies concern the prosodic marking of new information (Bortfeld & Morgan, 2010; Fernald & Mazzie, 1991; Fisher & Tokura, 1995), we specifically manipulate the familiarity of words to establish word-learning contexts in Chapters 2 and 3. Thus, the first research question in Chapters 2 and 3 is:

> **RQ 1**: What prosodic means do mothers use in IDS in word-learning contexts in which mothers introduce unfamiliar words to children?

Second, studies on the prosody of IDS often focus on a single language without cross-linguistic investigations of prosodically distinct languages. Thus, we examined the prosody of Dutch and Mandarin Chinese IDS in word-learning contexts using similar speech elicitation methods and materials. Thus, the second research question regards potential cross-linguistic differences in prosodic input in word learning.

RQ 2: How does the prosody of IDS in word-learning contexts (compared to ADS) differ between Dutch and Mandarin Chinese?

Third, most studies on IDS focus on the first year of life, while relatively fewer studies have examined IDS addressed to children during the “vocabulary spurt” period from 18 to 24 months. It remains unclear whether the global prosody of IDS is still exaggerated and whether it changes during this period as children’s vocabulary size rapidly increases. Also, since this period is crucial for word learning, it is possible that the prosody of IDS in word-learning contexts also changes with age during this period. Thus, the third research question concerns the age-related changes of IDS prosody between 18 months and 24 months:

RQ 3: Does the global prosody of IDS and IDS prosody in word-learning contexts change from 18 to 24 months of age?

Chapter 2 and Chapter 3 address these three questions with regard to different aspects of prosody. **Chapter 2** aims to examine the speaking rate of IDS in word-learning contexts by comparing Dutch and Mandarin Chinese. Specifically, we asked (1) whether Dutch and Chinese mothers specifically slow down to highlight unfamiliar words; (2) whether IDS is slower than ADS in Dutch and Mandarin Chinese; and (3) whether the speaking rate of IDS and IDS in word-learning contexts changes when mothers address children from 18 to 24 months. **Chapter 3** investigates pitch properties in word-learning contexts in Dutch and Mandarin IDS. In particular, we asked (1) whether mothers make distinctions between unfamiliar and familiar words with pitch in IDS compared to ADS; (2) how Mandarin Chinese and Dutch IDS differ in their pitch properties in word-learning contexts in which mothers introduce unfamiliar words to children; and (3) whether pitch properties of IDS and IDS in word-learning contexts change when mothers address children from 18 to 24 months.

Mandarin Chinese is a tonal language, in which lexical tones are crucial to distinguish word meanings. As such, it is important to examine lexical tones in IDS. However, whether lexical tones in Mandarin Chinese IDS are hyperarticulated is still in question from previous studies. Furthermore, as previous studies that tested tone hyperarticulation focused on the first year of age, it remains unknown whether lexical tones are still hyperarticulated in the second year when children start to learn words. **Chapter 4** reports a study on lexical tones in Mandarin Chinese IDS addressed to 18- and 24-month-old children. Specifically, this chapter investigates:

RQ 4a: Are tones in Mandarin Chinese IDS addressed to 18- and 24-month-old children hyperarticulated compared to tones in ADS?

RQ 4b: Does the hyperarticulation of lexical tones in Mandarin Chinese IDS change when mothers address an 18-month-old child versus a 24-month-old child?

As it remains unclear whether there is a correlation between the prosody of IDS and children's vocabulary size, we explored the link between the prosody of IDS and children's vocabulary size in **Chapter 5**. Specifically, we asked the following research questions:

RQ 5a: Are there correlations between the generally exaggerated prosody of IDS and children's vocabulary size?

RQ 5b: Are there correlations between the prosody of IDS specific in word-learning contexts and children's vocabulary size?

There is no sufficient evidence that prototypical IDS facilitates children's word-to-object mapping across languages. Also, it remains unclear whether there is a correlation between the prosody of IDS and children's online word-to-object mapping. Thus, **Chapter 6** has two research questions:

First, we extended Ma et al. (2011) to Dutch, asking:

RQ 6a: Does prototypical IDS facilitate Dutch 24-month-old children's word-to-object mapping compared to ADS?

Second, we further explored the link between individual IDS prosody and children's word learning performances in ADS and IDS. In particular, we asked:

RQ 6b: Does the degree of prosodic exaggeration in IDS predict children's performances in word-to-object mapping in ADS and IDS?

Chapter 7 includes a general discussion of all chapters, provides insights into future work, and concludes the whole dissertation.

An adapted version of Chapter 4 has been published in the journal *Frontiers in Psychology*. Partial results of Chapter 2 have been published in the *Proceedings of the 42nd annual Boston University Conference on Language Development*. Chapter 3 is under revision for publication in a journal. Adapted versions of Chapters 2, 5, and 6 will be submitted to appropriate journals.

Chapter 2

Speaking rate of infant-directed speech in word-learning contexts: A cross-linguistic investigation of Dutch and Mandarin Chinese*

Abstract

It is widely accepted that infant-directed speech (IDS) is slower than adult-directed speech (ADS) across languages, and that a slow speaking rate facilitates children's word learning. However, recent evidence suggests that a slow speaking rate may not be a common characteristic of IDS across languages or across ages. Also, no study to date has investigated whether mothers slow down their speaking rate in word-learning contexts in which they introduce unfamiliar words to children. This study uses a semi-spontaneous storybook-telling task to examine (1) whether IDS is slower than ADS in Dutch and Mandarin Chinese; (2) whether Dutch and Chinese mothers specifically slow down to highlight unfamiliar words; and (3) whether the speaking rate of IDS changes when addressing children from 18 to 24 months. Results suggest that Dutch IDS was slower than ADS, but Mandarin Chinese IDS did not show evidence of slowing down compared with ADS. In word-learning contexts, Dutch mothers slowed down specifically when introducing unfamiliar words in IDS compared to familiar words. In both languages, there was no evidence of age-related changes in speaking rate from 18 to 24 months. These results suggest that the temporal modifications in IDS, especially the effect of word familiarity on the speaking rate of IDS, may differ among languages. Consequently, the role of IDS in word learning and the specific prosodic cues that may account for the potential facilitative effects of IDS require further examination in a diversity of languages.

* Portions of this chapter have been published as: Han, M., de Jong, N. H., & Kager, R. (2018b). Infant-directed speech is not always slower: Cross-linguistic evidence from Dutch and Mandarin Chinese. In A. Bertolini and M. Kaplan (Ed.) *Proceedings of the 42nd annual Boston University Conference on Language Development* (pp. 331–344). Cascadilla Press.

2.1 Introduction

Infant-directed speech (IDS) is an important type of linguistic input in children's early language development across languages and cultures. Previous studies suggest that IDS has an exaggerated prosody compared to adult-directed speech (ADS). For example, prototypical IDS is characterized by a higher pitch, a larger pitch range, and a slower speaking rate compared to ADS (see a review in Soderstrom, 2007). Regarding speaking rate, the focus of this chapter, while it is widely assumed that IDS is universally slower than ADS (Martin et al., 2016), there is some recent evidence to suggest that a comparatively slower speaking rate of IDS might not hold universally. In particular, IDS is slower than ADS in some languages but not notably slower in others (Narayan & McDermott, 2016). IDS is not slower than ADS across entire utterances (e.g., Church, Bernhardt, Shi, & Pichora-Fuller, 2005; Martin, Igarashi, Jincho, & Mazuka, 2016). Moreover, there is some evidence to suggest that the speaking rate of IDS increases with children's age (Narayan & McDermott, 2016; Sjons et al., 2017). Existing studies on IDS speaking rate often focus on a single language and on one age group while diverging by adopting different speech elicitation methods and measures. As such, it is difficult to compare results directly from a cross-linguistic point of view. In light of the possible variations between languages, ages, and methods, we conducted a cross-linguistic investigation on the speaking rate of Dutch and Mandarin Chinese IDS using similar speech elicitation methods.

Slow speaking rate has been shown to facilitate children's online word learning compared to fast speech (Song, Demuth, & Morgan, 2010; Zangl & Mills, 2007), suggesting that the slow speaking rate of IDS has potential linguistic functions. However, whether the speaking rate in IDS is varied in a way that may support word learning is unclear. Specifically, no research has examined to what extent mothers vary speaking rate in word-learning contexts, in which they introduce unfamiliar words to children.

Speaking rate is influenced by many factors, including speech register, word frequency, word position in an utterance, word type, utterance length, and information status, to name a few (Quené, 2007; Seifart et al., 2018). The current study focuses on the effect of speech register (ADS or IDS) and the familiarity of words (familiar or unfamiliar).

2.1.1 Speaking rate of IDS across languages

Previous studies that have measured IDS speaking rate usually focus on a single language instead of examining different languages using similar speech elicitation methods. Some studies suggest that IDS is slower than ADS in the language under investigation. For example, Fernald and Simon (1984) showed that German IDS addressed to newborns (3 to 5 days old) had shorter utterances, longer pauses, and a slower articulation rate (excluding silent pauses and measured in syllables per second) compared with ADS. American English addressed to toddlers (17 to 20 months of age) was found to have a slower speech rate (including silent pauses and measured in words per minute) in comparison with ADS (Bernstein Ratner, 1985). In a small-scale study on Cantonese IDS ($N = 7$), Tang and Maidment (1996) showed that mothers' speech rate was slower in IDS directed towards 12- to 20-month-old children compared with ADS. The articulation rate of Swedish IDS addressed to 7- to 33-month-old children was also slower than ADS (Sjons, Hörberg, Östling, & Bjerva, 2017). Narayan and McDermott (2016) is to date the only known cross-linguistic study on IDS articulation rate. In this study, the articulation rates of Sri Lankan Tamil, Tagalog, and Korean IDS were compared with that of ADS in the three respective languages. The speech data were from natural mother-child interactions at home when children were 4 to 16 months of age. Their results demonstrated cross-linguistic differences: Tagalog and Korean mothers spoke slower in IDS addressing 4- to 16-month-old children compared to ADS, while Sri Lankan Tamil IDS was not slower than ADS for any of the infant age groups under investigation.

In addition to cross-linguistic differences, even in languages where IDS is slowed down globally compared to ADS, IDS is not necessarily slowed down across entire utterances. For instance, the articulation rate of Canadian English IDS addressed to preverbal children was slowed down in IDS when compared to ADS, but the difference between IDS and ADS disappeared when utterance-final syllables were excluded, suggesting that IDS is not slower in non-final syllables (Church, Bernhardt, Shi, & Pichora-Fuller, 2005). Similar results were found in Japanese IDS addressed to toddlers (17- to 25-month-olds), where IDS was only slowed down in phrase-final and utterance-final positions (Martin et al., 2016).

The inconsistent results in the studies reviewed above may be attributed to cross-linguistic differences, yet the mixed results could also have resulted from methodological differences between the studies. While the IDS and ADS speech samples were usually collected from natural mother-child interactions and mother-experimenter conversations, respectively, the content and contexts of these spontaneous speech data may have varied between ADS and IDS and may have also differed among the studies. Language input in different activities differs to a large degree (Tamis-LeMonda, Custode, Kuchirko, Escobar, & Lo, 2018), thus the mixed results may simply be due to the different speech contexts. In addition, two common measurements of speaking rate were used in the studies reviewed above: (1) speech rate, which combines the speed of speech with number and duration of silent pauses, and (2) articulation rate, which is the speed of speech only, exclusive of silent pauses. It is important to distinguish these two measurements because slower speech rate may be due to the number and duration of silent pauses in addition to a slower articulation rate. Consequently, though some studies suggest that IDS is slower in speech rate, it remains unknown whether the articulation rate in these studies was also slower.

In summary, there is still no conclusive evidence that the articulation rate of IDS is slower than ADS across languages or across entire utterances based on previous studies. Crucially, a cross-linguistic investigation of IDS speaking rate using similar speech elicitation methods is lacking. Thus, our study sets out to fill this gap by examining IDS speaking rate in two typologically distant languages—Dutch (a Germanic language) and Mandarin Chinese (a Sinitic language)—using similar speech elicitation methods. The two languages differ in several prosodic dimensions with respect to durational properties. For example, first, Dutch has lexical stress while Mandarin Chinese does not have lexical stress. In stress languages, lexical stress is placed on a given syllable in a word. The main acoustic correlates of Dutch lexical stress are duration and spectral balance (Sluijter & Van Heuven, 1996). Second, in terms of speech rhythm, there is some acoustic evidence to suggest that Dutch is closer to the typological extreme of stress-timed languages (e.g., Ramus, Dupoux, & Mehler, 2003), while Mandarin Chinese is closer to the typological extreme of syllable-timed language (e.g., Mok & Dellwo, 2008). In stress-timed languages, stressed and unstressed syllables are distinguished in terms of duration and syllable weight, while syllable-timed languages have nearly equal weight and time in all syllables. Also, the durational variability is greater in stress-timed languages compared to syllable-timed languages (Grabe and Low, 2002). The temporal differences at the word level and at the rhythmic level might interact with the generally temporal modifications in IDS. Thus, it is possible that the temporal modifications of Dutch and Mandarin Chinese IDS manifest themselves differently. Taken together, the first goal of the current study is to examine whether articulation rate is slower in IDS compared to ADS for both Dutch and Mandarin Chinese.

2.1.2 The effect of IDS speaking rate on word learning

If the speaking rate of IDS differs from ADS, a question that follows is: how might the differences in speaking rate, if any, influence language acquisition, and in particular, word learning? The relationship between IDS speaking rate and lexical development has been shown both concurrently and longitudinally. In a word recognition experiment, Song et al. (2010) manipulated the speech rate in English-learning children's auditory input: half of the stimuli were presented in typical English

IDS with a slow speech rate (1.94 syllables/s), while the other half had a faster speech rate (3.88 syllables/s) but preserved other prosodic aspects of IDS (such as higher pitch and larger pitch range). The results indicated that a slower speaking rate improved children's word recognition performance in comparison to faster speech. In another study, children's performance in word recognition was similarly found to improve when a word was presented in IDS articulation rate, which was twice as slow as ADS (Zangl, Klarman, Thal, Fernald, & Bates, 2005). Finally, a longitudinal study showed that slow speaking rate in IDS at seven months predicts larger expressive vocabulary at two years of age (Raneri, 2015).

Despite evidence from word recognition studies and a positive correlation between slow speaking rate and vocabulary size, it remains unknown whether mothers specifically vary their speaking rate when they introduce unfamiliar words to children compared to familiar words in order to promote word learning. Hyper and Hypo-speech (H&H) theory suggests that speakers are aware of the information required by a listener and adapt their speech accordingly (Lindblom, 1990). H&H theory was initially proposed to explain the phonetic variations in speech and has been often used to explain the "vowel hyperarticulation" phenomenon in IDS which may facilitate children's categorical learning (e.g., Kuhl et al., 1997). Fernald (2000) further proposed that H&H theory can be extended to explain how caregivers modify their speech to potentially facilitate children's word recognition. For example, both American English and Japanese IDS have more words produced in isolation and have more repetitions compared to ADS (Fernald & Morikawa, 1993). In English IDS, contextually new words (focused words) are often in utterance-final position (Fernald & Mazzie, 1991). Accordingly, the author proposed that "[…] the idea that adults intuitively and dynamically accommodate speech to infants in order to make their meanings more accessible to inexperienced listeners is still plausible." (Fernald, 2000, p. 250). In this chapter, we take a further step by examining whether IDS speaking rate is varied in a way that may support children's word learning. As children may benefit from slow speaking rate not just when recognizing words but also when learning new words, if speaking rate in IDS indeed facilitates word learning, it can be predicted that articulation rate in IDS would be specifically slowed down when mothers introduce words that children do not understand (unfamiliar words) compared to when introducing words that children understand (familiar words). That is to say, the familiarity of words may affect speaking rate in IDS such that mothers slow down more when introducing unfamiliar words compared to introducing familiar words. In sum, our second goal is to explore the variation of IDS speaking rate in word-learning contexts in which mothers introduce words that are unfamiliar to children.

The effect of familiarity on speaking rate in IDS may be potentially confounded with the effect of word frequency on speaking rate. As the age of acquisition of words is highly correlated with word frequency (Baker & Bradlow, 2009; Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012), words that are unfamiliar to children tend also to be less frequent in ADS compared to familiar words. In ADS, speakers slow down when the information conveyed is important, less predictable, or the words are less frequent or take longer to retrieve (e.g., Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Cohen Priva, 2017; Lieberman, 1963). It can thus be predicted that a mother's articulation rate in ADS would be affected by word frequency; her articulation rate would be slower in ADS when she uses less frequent words compared to frequent words. These differences in articulation rate, however, would then be magnified in IDS when the effect of familiarity is added.

2.1.3 Age-related changes

In addition to the potential cross-linguistic differences and variations related to the familiarity of words, a child's age may also affect an adult's speaking rate in IDS. Pitch properties of IDS are known to change to adapt to different stages of language development. In the first year of life, infants undergo "perceptual reorganization" during which time their perception of phonetic categories shifts from language-universal to language-specific (Werker & Tees, 1984). At the same time, infants begin recognizing common words as early as 6 to 9 months (Bergelson & Swingley, 2012), they utter their first words around their first birthday (Bloom, 2002), and their vocabulary size increases at a rapid speed from 16 to 18 months onwards (Goldfield & Reznick, 1990). In general, IDS becomes more similar to ADS as children grow older and gain better linguistic competence (e.g., Kitamura et al., 2002). Could this trend be extended to speaking rate? That is to say, does IDS speaking rate increase over time as children grow older as well?

So far only a few studies have explored age-related changes in IDS speaking rate. By examining Swedish IDS in a longitudinal corpus, Sjons et al. (2017) found that the articulation rate of IDS slightly increased from 7 months to 33 months, but was always slower than ADS. In their longitudinal investigation of speaking rate from 4- to 16-month-old children, Narayan and McDermott (2016) showed that overall, the speaking rate of IDS in Korean, Tagalog, and Sri Lankan Tamil increased with age, and was similar to ADS at 16 months in all three languages. As the age groups under investigation were undergoing perceptual reorganization, the authors suggested that the overall increase of speaking rate across time indicated that the properties of IDS might coincide with children's increasing ability in their native-language perception. Nevertheless, it is difficult to draw any conclusions regarding the developmental path of IDS speaking rate based on the small number of studies on different languages. The third goal of the current study, therefore, is to investigate age-related changes in IDS speaking rate. Since we specifically target IDS speaking rate in relation to word learning, we chose a period when children are having a vocabulary spurt from 18 months to 24 months.

2.1.4 The current study

To summarize, we had three research questions: (1) Is IDS addressed to 18- and 24-month-old children slower than ADS in Dutch and Mandarin Chinese? We predicted that IDS would be slower than ADS in both languages, consistent with previous studies on typologically related languages (Dutch and German or Mandarin Chinese and Cantonese). (2) Do mothers specifically slow down when introducing unfamiliar words compared to familiar words in Dutch and Mandarin Chinese IDS? We predicted that mothers would slow down for unfamiliar words compared with familiar words in both ADS and IDS because unfamiliar words are less frequent than familiar words in ADS, but the degree of slowing down would be larger in IDS compared with ADS condition due to mothers' highlighting of unfamiliar words to facilitate word learning. (3) Does the global speaking rate in IDS and IDS speaking rate specific in word-learning contexts change from 18 months to 24 months, during the vocabulary spurt period? We predicted that the global IDS speaking rate would become faster from 18 months to 24 months as children's vocabulary is rapidly increasing during this period. With respect to the speaking rate of IDS specific to word-learning contexts, it is possible that it will become faster from 18 months to 24 months, consistent with the global speaking rate. Alternatively, it may show similar patterns at these two ages as mothers may keep slowing down unfamiliar words during this period.

To address the three research questions, we conducted two experiments on Dutch (Experiment 1) and Mandarin Chinese (Experiment 2) using similar materials and speech elicitation methods. The major difference between the two experiments was that we adopted a longitudinal design in the Dutch experiment and used a cross-sectional design in the Mandarin Chinese experiment.[1] We measured the articulation rate of words and utterances in IDS in a semi-spontaneous storybook-telling task. In addition, we explored the role of word position by examining whether mothers specifically put unfamiliar words significantly more frequently at utterance-final positions than familiar words in IDS.

---

[1] The difference in design was mainly due to the practical situation in which we recruited our participants in China. The participants were mostly recruited from early education programs in kindergartens where they did not enroll for longer than a semester (6 months).

2.2 Methods

2.2.1 Experiment 1: Dutch

2.2.1.1 Participants

Thirty-two Dutch-speaking mother-child dyads participated when children were 18 months old (mean age of children = 18;15, age range = 18;00–18;29; girls *N* = 18; mean age of mothers = 35 years, age range = 29–44 years). All mothers had higher education.[2] The same participants visited the lab again when the children reached 24 months old (mean age of children = 24;18, age range = 24;00–26;30). The Dutch mother-child dyads were recruited from the Utrecht Baby Lab database and were all Dutch native speakers living in the Utrecht area in the Netherlands. All children were typically developing and none of them had any hearing problems or known developmental delays.

2.2.1.2 Materials

For the Dutch 18-month-old and 24-month-old children, two picture books were designed to elicit two sets of seven target words, with five unfamiliar words and two familiar words in each set (see **Table 2.1** for a list of target words). The book structure was the same for the 18-month-old and 24-month-old group, however, the five unfamiliar words were replaced with new unfamiliar words in the 24-month-old version. On each page of the picture book, a word was on the left side and an illustration including a depiction of the word was shown on the right side. No other script was provided besides the target words (see **Appendix A** for the picture book). An additional six pages were used as fillers to make the story coherent throughout the book.

[2] HBO (*hogescholen*, "universities of applied sciences") or WO (*universiteiten*, "research universities") and above.

Table 2.1. Target words in Experiment 1 and Experiment 2.

Default Familiarity	Dutch 18 months old	Chinese 18 and 24 months old (pinyin)	English translation	Dutch 24 months old	English translation
Familiar	opa	yé ye	"grandpa"	opa	"grandpa"
Familiar	appel	píng guǒ	"apple"	appel	"apple"
Unfamiliar	eland	mí lù	"moose"	emoe	"emu"
Unfamiliar	bever	hé lí	"beaver"	wezel	"weasel"
Unfamiliar	walnoot	hé tao	"walnut"	bamboe	"bamboo"
Unfamiliar	kasteel	chéng bǎo	"castle"	kapel	"chapel"
Unfamiliar	pompoen	nán guā	"pumpkin"	jasmijn	"jasmine"

The target words were all disyllabic nouns. As we intended to use similar experimental materials in each language, we selected familiar words that were both listed in the Dutch (N-CDI, Zink & Lejaegere, 2002) and Mandarin Chinese (M-CDI, Tardif, Fletcher, Liang, & Kaciroti, 2009) versions of MacArthur-Bates Communicative Development Inventories (CDI) (Fenson, Bates, Dale, Marchman, & Reznick, 2007). The unfamiliar words were not listed in N-CDI or M-CDI. Also, the familiar words were more frequent than the unfamiliar words in each language.[3] Selecting target words in such a way ensured that the default familiarity of the words applied to most of the participants. However, due to individual differences in vocabulary knowledge, the actual familiarity of the target words might vary among children. To examine whether each child was familiar with the target words or not, mothers filled out a word checklist after the experiment to determine whether their child had understood the target words before reading the picture book. This information was coded as Familiarity (Familiar/Unfamiliar) and used in data analyses.

[3] The ranking (lower rank indicating a higher frequency) of Mandarin Chinese word frequency based on Cai and Brysbaert (2010) is: *yé ye* ("grandpa") (1662), *píng guǒ* ("apple") (2939), *mí lù* ("moose") (17914), *hé lí* ("beaver") (55578), *hé tao* ("walnut") (12883), *chéng bǎo* ("castle") (3149), and *nán guā* ("pumpkin") (5744). The ranking of Dutch word frequency according to Keuleers, Brysbaert, and New (2010) is: *opa* ("grandpa") (1211), *appel* ("apple") (4666), *eland* ("moose") (12385), *bever* ("beaver") (11515), *walnoot* ("walnut") (28953), *kasteel* ("castle") (2185), *pompoen* ("pumpkin") (12830), *bamboe* ("bamboo") (30072), *wezel* ("weasel") (14576), *emoe* ("emu") (76161), *kapel* ("chapel") (8604), *jasmijn* ("jasmine") (26190). Note that word frequency is only provided to show that unfamiliar words overall have a lower word frequency. Ranking is not comparable between languages. We used the mothers' reports as an indication for Familiarity in analyses.

2.2.1.3 Procedure

Each participant came to the lab twice at 18 months and 24 months. All participants were tested in a quiet room in the Utrecht Baby Lab. Before the experiment, mothers were given a few minutes to familiarize themselves with the book. Each experiment consisted of two conditions: IDS condition and ADS condition. In the IDS condition, the child sat on his or her mother's lap, and the mother was instructed to tell the story to her child the way she usually would at home. The mothers were specifically told that they could use any sentences; the only requirement was to include the words given on each page. In the ADS condition, the mothers were instructed to tell the story to the experimenter (female, a native speaker of Dutch), and to take into account the fact that she was a college student. The order of the two conditions was counterbalanced across participants. A ZOOM H1 recorder (with 16-bit resolution and a sampling rate of 44.1 kHz) was used to make audio recordings. Each experimental session took about 15–20 minutes. All families received a book as a gift after the experiment.

2.2.2 Experiment 2: Mandarin Chinese

2.2.2.1 Participants

Twenty-one Mandarin-Chinese-speaking[4] mothers of 18-month-old children (mean age = 18;15, age range = 17;21–18;27; girls $N = 9$; mean age of mothers = 30 years, age range = 25–39 years) and nineteen mothers of 24-month-old children (mean age = 24;13, age range = 23;27–24;30; girls $N = 10$; mean age of mothers = 31 years, age range = 32–36 years) participated in the study. All mothers had higher education.[5] The Mandarin Chinese dyads were recruited from kindergartens in Yichang, China. All the participant mothers spoke Mandarin Chinese (the official language in China) proficiently and they spoke Standard Mandarin to their children on a regular basis.[6] All children were typically developing and none of them had any hearing problems or known developmental delays.

---

[4] We use the term "Mandarin Chinese" in reference to "Putonghua" or "Standard Chinese", the official language spoken in China. It should be distinguished from Taiwanese Mandarin, another variety of Mandarin Chinese spoken in Taiwan.

[5] Undergraduate degree and above.

[6] All the participant mothers spoke Mandarin Chinese and a dialect (Southwest Mandarin). The participant children heard this dialect in their language community, but were exposed to Mandarin Chinese at home, at kindergarten, and in the national media. This type of bilingual language background is common for most people in China (Li & Lee, 2006). We set these criteria in our recruiting interview: (1) the mothers should speak Mandarin Chinese with good proficiency; (2) the mothers should mostly speak Mandarin Chinese to their children at home; and (3) the children should be learning Mandarin Chinese as one of their first languages.

2.2.2.2 Materials and procedure

The picture book for the Dutch 18-month-old group was adapted to Mandarin Chinese and was used for the Mandarin Chinese 18-month-old and 24-month-old children (see **Appendix A** for example pages of the picture book). The Mandarin Chinese participants were tested in a quiet room. The procedure was identical to Experiment 1, however the experimenter was a native Mandarin Chinese speaker (female).

2.2.3 Data analysis

A trained Dutch native speaker and a Chinese native speaker (the author) annotated and extracted the target words and target utterances (utterances containing the target words) from the recordings using Praat (Boersma & Weenink, 2017). An utterance boundary was defined in accordance with Martin et al. (2016): "any pause longer than 200ms which is preceded by an intonational phrase boundary (pauses not accompanied by an IP boundary were considered utterance internal)" (Martin et al., 2016, p. 54). Silent pauses mostly occurred between utterances, and sometimes within utterances. The durations of silent pauses were subtracted from the utterance durations when calculating the articulation rate of utterances. The annotators also transcribed and manually counted the numbers of phonological syllables for the target utterances. In total, 1521 Dutch utterances and 1375 Chinese utterances were included for further analysis, among which there were 579 utterances with familiar words in Dutch (ADS: 244) and 857 utterances with familiar words in Mandarin Chinese (ADS: 335). The total duration of the speech sample was 44.13 minutes for Dutch (ADS: 21.61 min; IDS: 22.52 min) and 48.13 minutes for Mandarin Chinese (ADS: 22.2 min; IDS: 25.93 min). We measured the articulation rate of the target words as well as the target utterances (see **Table 2.2** for the formulas of each measurement).

Table 2.2. Measurements on target utterances and formulas.

Measurements on target words and utterances	**Formulas**
Target word articulation rate (syllables/s)	Number of syllables of words/Target word duration
Target utterance articulation rate (syllables/s)	Number of syllables of utterances/ (Utterance duration – pause duration)

In the models, we included fixed factors of Age (18 months/24 months), Condition (ADS/IDS), and Familiarity (Familiar/Unfamiliar) with Participant as a random factor. In Dutch, we allowed for random slopes for Age, Condition, and Familiarity (Barr, Levy, Scheepers, & Tily, 2013). In Mandarin Chinese, due to the cross-sectional design, we included Condition and Familiarity as random slopes but not Age. The dependent variables were target word articulation rate and target utterance articulation rate. The dependent variable target word articulation rate for Mandarin Chinese was log-transformed to approximate a normal distribution.

We used the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in the R environment for data analyses (R Core Team, 2018). For each dependent measure, we took the backward elimination approach, starting with a model that included all fixed effects and all interactions between them, plus the random factor (the most complex model[7]) (Bates, Kliegl, Vasishth, & Baayen, 2015). Then, we used the "step" function in the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) to reduce the models by eliminating non-significant factors or interactions. The final models are reported in the paper, and the means and standard deviations are presented in **Table 2.3**.

2.3 Results

2.3.1 Experiment 1: Dutch

We checked whether there was an effect of testing order (ADS-IDS/IDS-ADS) and no significant differences were found between the two testing orders for any of the dependent measures. With respect to the research questions, first, are unfamiliar words specifically slowed down in IDS compared with ADS? The box plots of word articulation rate in Dutch ADS and IDS at 18 months and 24 months are presented in **Figure 2.1**. Our final model for Dutch target word articulation rate revealed that there were no significant interactions, but there was a significant main effect of Condition ($p < 0.001$) and a significant main effect of Familiarity ($p < 0.001$) (**Table 2.4a**). These results suggest that both familiar and unfamiliar words were slower in IDS than in ADS, confirming our prediction that Dutch IDS would be slower than ADS. Also, unfamiliar words were slower than familiar words in both ADS and IDS, which could be interpreted as an effect of word frequency on articulation rate. The model also showed a significant main effect of Age ($p < 0.001$), suggesting that word articulation rate at 24 months was faster than at 18 months. As there was no interaction with Condition or Familiarity, this age effect might be attributed to the different unfamiliar words selected at the two different ages of testing. In sum, results showed that the target words were slower in Dutch IDS than in ADS, and the unfamiliar words were slower than familiar words. However, the degree of slowing down in IDS did not differ between familiar and unfamiliar words.

---

[7] An example of the R codes is: lmer(articulation_rate ~ Age * Condition * Familiarity + (1 + Age + Condition + Familiarity | Participant))

Table 2.3. Mean articulation rate and mean target word articulation rate in Dutch and Mandarin Chinese (standard deviations in parentheses).

Language	Familiarity	Condition	Target word articulation rate	Utterance articulation rate
Dutch 18 months	Familiar	ADS	6.01 (1.58)	5.15 (1.16)
		IDS	5.39 (1.74)	4.92 (1.18)
	Unfamiliar	ADS	4.58 (1.31)	5.12 (1.12)
		IDS	4.05 (1.26)	4.56 (1.28)
Dutch 24 months	Familiar	ADS	6.11 (1.66)	5.38 (1.11)
		IDS	5.67 (1.71)	5.09 (1.24)
	Unfamiliar	ADS	5.10 (1.27)	5.17 (0.97)
		IDS	4.37 (1.19)	4.68 (1.16)
Chinese 18 months	Familiar	ADS	4.61 (1.94)	4.56 (1.01)
		IDS	4.36 (1.91)	4.50 (1.44)
	Unfamiliar	ADS	4.04 (1.45)	4.28 (0.97)
		IDS	4.03 (2.03)	4.33 (1.38)
Chinese 24 months	Familiar	ADS	4.19 (1.61)	4.59 (1.05)
		IDS	4.46 (1.87)	4.84 (1.49)
	Unfamiliar	ADS	4.20 (1.31)	4.61 (1.01)
		IDS	4.04 (1.87)	4.37 (1.54)

If the unfamiliar words were not particularly slowed down in IDS compared with familiar words, how about the utterances containing unfamiliar words (see **Figure 2.2** for box plots of Dutch utterance articulation rate)? The final model for Dutch utterance articulation rate (**Table 2.4b**) showed that there was a significant main effect of Condition ($p = 0.008$) as well as a significant interaction of Condition and Familiarity ($p = 0.022$). These results indicated that at the utterance level, the articulation rate was slowed down in IDS, however the degree of slowing down was larger for utterances containing unfamiliar words in comparison with utterances containing familiar words. As the factor "Age" was not in the final model, these results held for both age groups, suggesting that there was no evidence of age-related changes in the articulation rate at the utterance level.

Figure 2.1. Box plots of word articulation rate (syllables/s) for ADS and IDS in Dutch.[8]

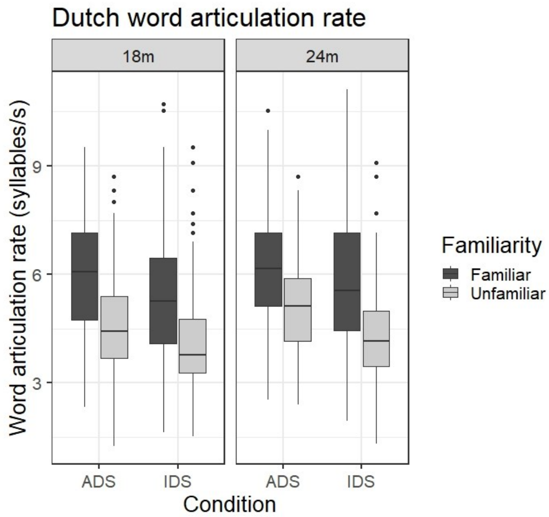
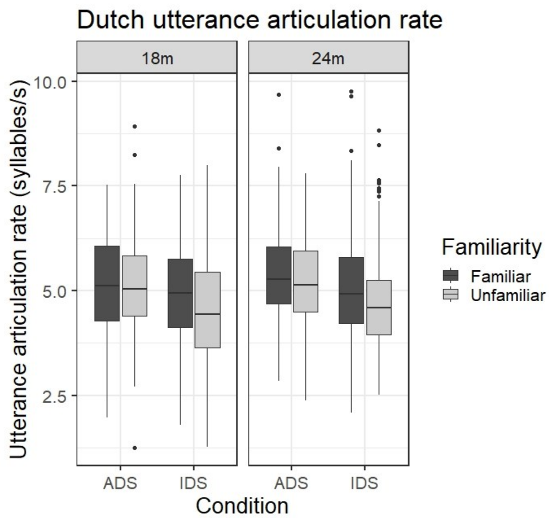


Figure 2.2 Box plots of utterance articulation rate (syllables/s) for ADS and IDS in Dutch.



[8] For all the box plots in this dissertation: The box area shows the 1st quantile (lower hinge) and 3rd quantile (upper hinge). The lower whisker indicates the smallest observation greater than or equal to lower hinge – 1.5 * IQR. The upper whisker shows the largest observation less than or equal to upper hinge + 1.5 * IQR. Each line within boxes denotes medians. Outliers are represented by dots.

Table 2.4. The final models for Dutch target word articulation rate (Table 2.4a) and utterance articulation rate (Table 2.4b).

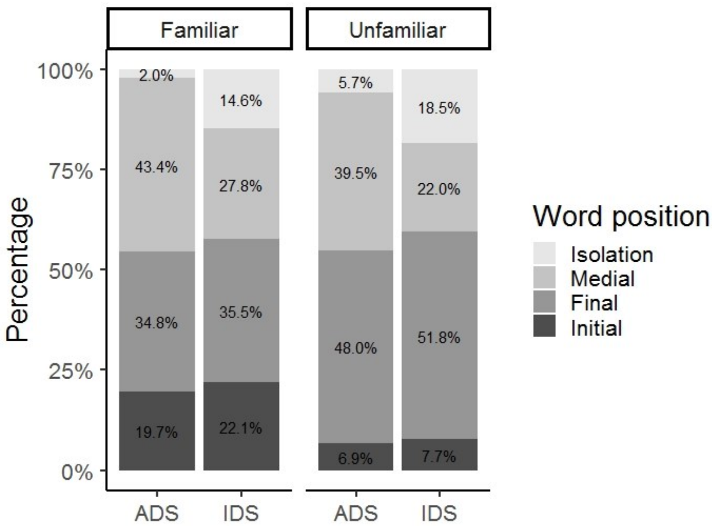
Table 2.4a. Final model for Dutch target word articulation rate				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	5.954	0.127	46.905	<0.001***
Age (24m)	0.316	0.071	4.439	<0.001***
Condition (IDS)	-0.590	0.071	-8.321	<0.001***
Familiarity(Unfamiliar)	-1.347	0.096	-14.053	<0.001***
	Random factors			
	Variance	***SD***		
Participant(Intercept)	0.315	0.561		
Familiarity (Unfamiliar)	0.120	0.347		
Residual	1.859	1.364		

Table 2.4b. Final model for Dutch utterance articulation rate				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	5.242	0.098	53.259	<0.001***
Condition (IDS)	-0.246	0.093	-2.649	0.008**
Familiarity(Unfamiliar)	-0.143	0.089	-1.595	0.111
Condition (IDS): Familiarity (Unfamiliar)	-0.270	0.118	-2.290	0.022*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	0.148	0.385		
Residual	1.212	1.101		

Note: For Table 2.4a, Intercept represents ADS, 18 months, and Familiar. For Table 2.4b, Intercept represents ADS and Familiar. $*\ p < 0.05$; $**\ p < 0.01$; $***\ p < 0.001$

As Martin et al. (2016) showed that final-lengthening contributed to the slower speaking rate in IDS, we also explored whether mothers put unfamiliar words significantly more frequently at utterance-final positions than familiar words in IDS. As there were no significant interactions between Age and Condition in the previous analyses, we collapsed the two ages for this analysis. **Figure 2.3** presents the distribution of target word positions in ADS and IDS for familiar and unfamiliar words. We performed a multinomial logistic regression using the multinom() function in the nnet package in R (Venables & Ripley, 2002). The outcome variable word position contained four levels: initial, medial, final, and isolation. We used the level "initial" position as the intercept.[9] The predictors were Condition (ADS/IDS) and Familiarity (Familiar/Unfamiliar). We then calculated the p-values using the Wald tests. The results showed that there was a significant effect of Condition on isolation position ($\beta = 1.849$, $SE = 0.505$, $p < 0.001$), indicating that target words occurred more frequently in isolation in IDS compared to ADS. There was also a significant effect of Condition on medial position ($\beta = -0.564$, $SE = 0.234$, $p = 0.016$), suggesting that target words occurred less frequently in medial position in IDS than in ADS. Crucially, there was no significant effect of Condition on final position, indicating that Dutch mothers do not put words at utterance-final positions more frequently in IDS compared to ADS. Also, there were no significant interactions of Condition and Familiarity on any of the position levels, suggesting that these results hold for both familiar and unfamiliar words.

Figure 2.3 Word position of the target words in Dutch.



To summarize, our Dutch results showed that the articulation rate of the target words and utterances with target words were consistently slower in IDS compared with ADS. However, utterances with unfamiliar words were slowed down to a greater

[9] An example of the R codes:

```
lmtest <- multinom(formula = Word_position ~ Condition * Familiarity, data)
z <- summary(lmtest)$coefficients/summary(lmtest)$standard.errors
p <- (1 - pnorm(abs(z), 0, 1)) *2
```

extent in IDS compared to utterances with familiar words. No age-related changes were found in the articulation rate of IDS at either word or utterance level.

2.3.2 Experiment 2: Mandarin Chinese

The box plots of word articulation rate in Mandarin Chinese ADS and IDS at 18 months and 24 months are presented in **Figure 2.4**. We performed similar analyses on the Mandarin Chinese data as on the Dutch data. The dependent measure word articulation rate was log- transformed from raw data to get a more normalized distribution. Condition was not in the final model for word articulation rate (**Table 2.5a**), suggesting that the target words did not show evidence of slowing down in IDS regardless of Familiarity. Unfamiliar words were slower than familiar words in both ADS and IDS, shown by a significant main effect of Familiarity (p = 0.017). The significant main effect of Familiarity may be explained by the word frequency effect. These results held for both age groups because Age was not in the final models, suggesting that there was no evidence of age-related changes in word articulation rate from 18 months to 24 months. For the measure of utterance articulation rate (**Table 2.5b**), we excluded two outliers that were more than 3 standard deviations from the mean. The box plots of utterance articulation rate in Mandarin Chinese ADS and IDS at 18 months and 24 months are presented in **Figure 2.5**. Results at the utterance level were similar to the results at the word level: there was no effect of Condition, but a significant main effect of Familiarity (p = 0.036), indicating that utterances with unfamiliar words were slower than utterances with familiar words in both ADS and IDS.

Figure 2.4. Box plots of word articulation rate (syllables/s) for ADS and IDS in Mandarin Chinese (y-axis is log-transformed).

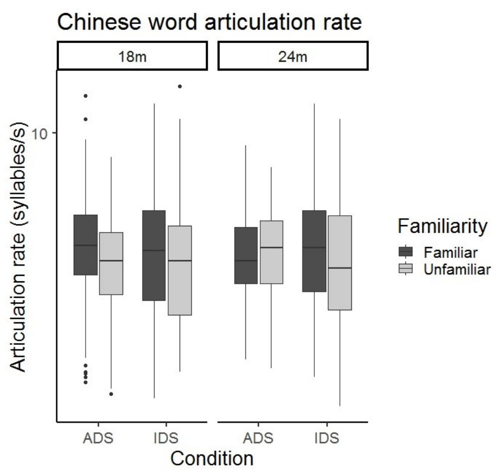


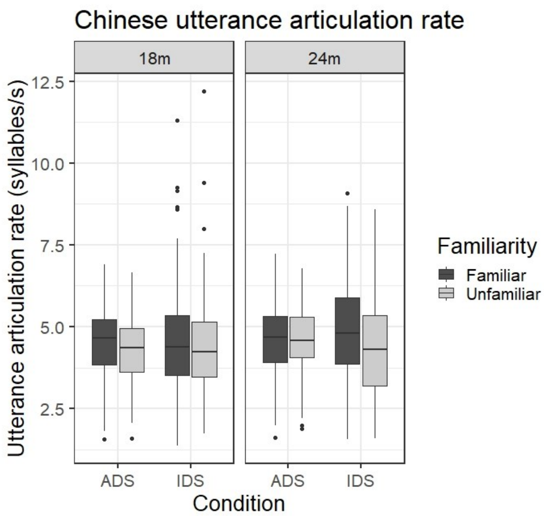
Table 2.5. The final models for Mandarin Chinese target word articulation rate (log-transformed) (Table 2.5a) and utterance articulation rate (Table 2.5b).

Table 2.5a. Final model for Mandarin Chinese target word articulation rate				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	1.383	0.024	51.437	<0.001***
Familiarity (Unfamiliar)	-0.056	0.023	-2.401	0.017*
	Random factors			
	Variance	***SD***		
Participant (Intercept)	0.024	0.156		
Condition (IDS)	0.029	0.170		
Residual	0.150	0.387		

Table 2.5b. Final model for Mandarin Chinese utterance articulation rate				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	4.600	0.092	49.865	<0.001***
Familiarity (Unfamiliar)	-0.211	0.096	-2.196	0.036*
	Random factors			
	Variance	***SD***		
Participant (Intercept)	0.287	0.536		
Condition (IDS)	0.328	0.573		
Familiarity (Unfamiliar)	0.148	0.385		
Residual	1.347	1.161		

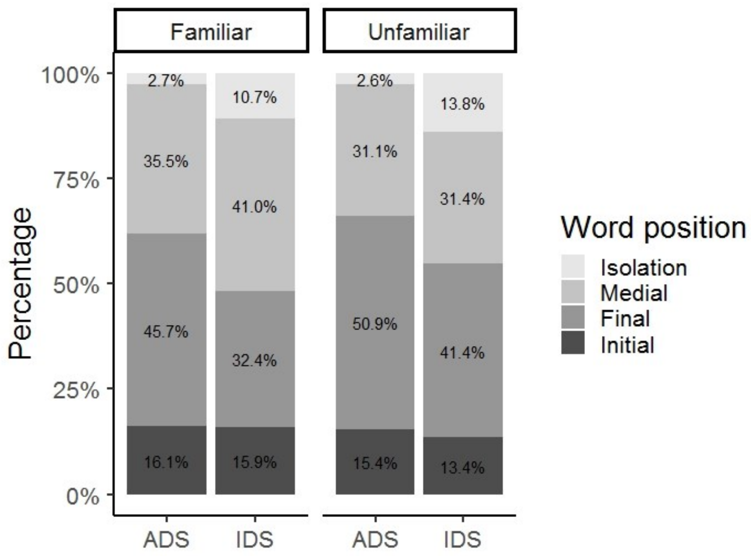
Note: Intercept represents Familiar in both Table 2.5a and Table 2.5b $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Figure 2.5. Box plots of utterance articulation rate (syllables/s) for ADS and IDS in Mandarin Chinese (two outliers are removed from analyses).



As in Dutch, we also performed a multinomial logistic regression to examine whether mothers prefer putting unfamiliar words in utterance-final position in IDS. **Figure 2.6** presents the distribution of target word positions in ADS and IDS for familiar and unfamiliar words. The initial position was used as intercept. The results showed that there was a significant effect of Condition on isolation position ($\beta = 1.398$, $SE = 0.399$, $p < 0.001$), indicating that the target words occurred more often in isolation in IDS compared to ADS. As in Dutch, there was no significant effect of Condition on final position, nor were there any significant interactions of Condition and Familiarity on any of the levels. These results suggest that mothers tend to put more target words in isolation but do not put more target words in utterance-final position in Mandarin Chinese IDS, and this trend holds for both familiar and unfamiliar words.

Figure 2.6. Word position of the target words in Mandarin Chinese.



Together these results suggest that there is no significant difference in articulation rate between Mandarin Chinese ADS and IDS. Mandarin Chinese mothers do not show evidence of slowing down their articulation rate in IDS when addressing 18-month-olds or 24-month-olds, even when they talk about words unfamiliar to the children. Similar to the Dutch results, Mandarin Chinese mothers put the target words more often in isolation in IDS compared to ADS, but they did not put target words more in final position in IDS compared to ADS.

2.4 Discussion

The current study focused on the temporal measure of IDS—speaking rate—and conducted a cross-linguistic investigation of Dutch and Mandarin Chinese using similar speech elicitation methods (a semi-spontaneous storybook-telling task). The purpose of the current study was to examine whether IDS was slower than ADS across languages and ages and whether the speaking rate in IDS was varied in a way that may potentially support word learning.

Our first research question asked whether IDS addressing 18- and 24-month-old children was slower than ADS in Dutch and Mandarin Chinese. The results showed that Dutch IDS, in accordance with previous studies on English and German, had a slower articulation rate than ADS (Bernstein Ratner, 1985; Fernald & Simon, 1984). However, the articulation rate of Mandarin Chinese IDS did not show evidence of slowing down at either word or utterance level. Cantonese, which is typologically related to Mandarin Chinese, was found to be slower in IDS than in ADS with respect to speech rate (Tang & Maidment, 1996). However, as speech rate includes silent pauses, it is unclear if the articulation rate in Cantonese IDS was slower than ADS or not. The Chinese results were in accordance with the results from Sri Lankan Tamil, which did not show evidence of slowing down its articulation rate in IDS (Narayan & McDermont, 2016). Second, we asked whether mothers slowed down when introducing unfamiliar words in particular compared to familiar words in IDS, and the Dutch results were consistent with our expectations. We found that Dutch mothers slowed down utterance articulation rate for unfamiliar words compared to familiar words in both ADS and IDS, but the degree of slowing down was larger in IDS compared to ADS, possibly due to mothers' highlighting of unfamiliar words. In contrast, Mandarin Chinese mothers did not show any evidence of slowing down even when they were introducing unfamiliar words, neither at the word nor at the utterance level. Third, does IDS speaking rate change from 18 months to 24 months during a vocabulary spurt? Our prediction was that global IDS speaking rate would become faster from 18 months to 24 months as children's vocabulary increases at a fast speed. Also, the speaking rate of IDS in word-learning contexts may or may not change with the global speaking rate. Our findings suggest that in both languages, the speaking rate of IDS as well as IDS specific to word-learning contexts did not show evidence of age-related changes from 18 to 24 months, regardless of whether IDS was slower or not.

A natural question arises as to why Dutch IDS was slower than ADS but Mandarin Chinese IDS did not show evidence of slowing down even though the speech elicitation method was similar for both languages. First, do these results suggest that the Mandarin Chinese mothers did not modify their prosody at all when talking to children? Although the main purpose of this chapter was to examine the temporal measure of the speech data, we compared the pitch height (mean F0) between ADS and IDS in Mandarin Chinese to rule out this possibility. By adopting analyses similar to those described in the methods section, we found that Mandarin Chinese IDS addressing 18-month-old children had a higher pitch than ADS though there were no differences in mean F0 between ADS and IDS addressing 24-month-old children (see Chapter 3 in this dissertation for more details on the pitch measures).[10] These results suggest that Mandarin Chinese mothers did modify pitch when addressing children, yet the articulation rate did not seem to slow down in IDS. In any case, we can rule out the possibility that Mandarin Chinese mothers did not produce IDS in our task. Another possibility could be attributed to the typological differences between the two languages. The word prosody and rhythm of the two languages differ. As a syllable-timed language (without lexical stress), Mandarin Chinese has nearly equal weight and time in all syllables, while in Dutch, a stress-timed language (with lexical stress), the stressed and unstressed syllables are distinguished qua syllable weight and duration at the word level. Grabe and Low (2002) suggest that the durational variability is greater in stress-timed languages compared to syllable-timed languages. As a result, rhythmic class may have an effect on the temporal modifications in IDS. However, there has been little empirical evidence of how speech rhythm affects temporal modifications in IDS, especially for syllable-timed languages. Only one known study has reported that IDS in Australian English, a stress-timed language with lexical stress, is slower and more rhythmic than ADS (Leong, Kalashnikova, Burnham, & Goswami, 2017). It is possible that as in Australian English, the rhythmic variations of Dutch are magnified in IDS, resulting in an overall slower articulation rate. On the other hand, Mandarin Chinese syllable duration may or may not be sensitive to the temporal modifications in IDS as syllable-timed Mandarin Chinese has less temporal variations compared to languages such as Dutch and Australian English.

Another alternative explanation for the different results from Dutch and Mandarin Chinese involves the age of the infants under investigation. Two studies have demonstrated that speaking rate in IDS changes in accordance with children's language development (Sjons et al., 2017; Narayan & McDermott, 2016). For example, Narayan and McDermott (2016) showed that the articulation rate of Tagalog

---

[10] 18 months: ADS: mean F0 = 237.48 (Hz), *SD* = 39.47; IDS: mean F0 = 260.52 (Hz), *SD* = 47.01; $p < 0.001$); 24 months: ADS: mean F0 = 250.36 (Hz), *SD* = 46.80; IDS: mean F0 = 256.43 (Hz), *SD* = 50.07.

and Korean IDS increased from 4 months onwards and was already similar to ADS as children reached 15-16 months (though individual variation exists). Swedish IDS slightly increased articulation rate from 7 months to 33 months, though it was still slower than ADS at 33 months (Sjons et al., 2017). As our participants were 18-month-old and 24-month-old children, it is possible that Mandarin Chinese IDS addressed to younger children is slower than ADS, but the articulation rate already becomes ADS-like when children are 18 months of age. One direction that follows from the current research is to examine the developmental trajectory of speaking rate in Mandarin Chinese IDS before 18 months old.

The developmental changes in speaking rate seem to differ across languages. Specifically, Dutch and Swedish IDS are both slower than ADS beyond two years of age. However, the articulation rate of Tagalog and Korean IDS becomes similar to ADS before 18 months (Narayan & McDermont, 2016). Also, Mandarin Chinese IDS does not show evidence of slowing down from 18 to 24 months of age. Taking evidence from these languages into account, it seems that IDS in western languages prolong the slower speaking style when addressing children, while languages and cultures such as Mandarin Chinese, Korean, and Tamil (all Asian cultures), drop the temporal modifications as early as 16 months. Future studies should examine the sources of the different developmental traceries of IDS speaking rate across languages.

Our results did not show evidence that Mandarin Chinese mothers vary speaking rate depending on the familiarity of words. However, this does not mean that Mandarin Chinese IDS does not differentiate familiar and unfamiliar words at all. In addition to the rhythmic differences between Dutch and Mandarin Chinese, another crucial difference between the two languages is the use of pitch at a lexical level: Mandarin Chinese is a tone language in which pitch is used to distinguish lexical meaning, while Dutch mainly uses pitch at the intonational level. One possibility is that Mandarin Chinese uses pitch-related cues such as pitch height and pitch range instead to mark unfamiliar words (see Chapter 3 in this dissertation).

Our findings extend past research on the cross-linguistic differences of IDS. Prior research has demonstrated that IDS prosody has exaggerated prosody across all languages, as the general prosodic modifications such as higher pitch and a larger pitch range have been found in a variety of languages. As a result, more often than not studies assume “only slight differences across languages and cultures” in IDS (e.g., Spinelli et al., 2017, p. 2). However, a majority of the studies on IDS are conducted on stress-timed languages or stress languages such as English, thus their results may not be applied to languages with different prosodic characteristics (Wang, Seidl, & Cristia, 2016). Even the few studies that paid attention to the cross-linguistic differences in IDS often aimed at examining whether the prosodic modifications were present in a specific language, or whether the degree of prosodic exaggeration in IDS was more prominent in some languages than others. For example, American English IDS was more exaggerated than British English, French, Italian, German, and Japanese IDS (Fernald et al., 1989); Thai IDS was less exaggerated than Australian English IDS (Kitamura et al., 2002). Our findings suggest that either the temporal modifications in IDS differ between Dutch and Mandarin Chinese, or familiarity affects the variation of speaking rate in Dutch and Mandarin Chinese IDS differently. These findings indicate that the cross-linguistic differences in IDS are not restricted to how exaggerated the prosodic modifications are. Instead, languages may employ different means to highlight unfamiliar words in IDS, which may in turn influence children's word learning strategies in meaningful ways. Consequently, the correlation between speaking rate and vocabulary size, and the facilitating effects of speaking rate on word recognition (Raneri, 2015; Song et al., 2010), may not be extended to languages in which speaking rate is not varied in a way that may support lexical development. Whether Mandarin Chinese children's lexical development benefits or suffers from slow speaking rate requires further investigation.

In addition to the cross-linguistic variations in IDS temporal modification, our study also demonstrated cross-linguistic differences in word order modifications in IDS. Fernald and Mazzie (1991) reported that English IDS tends to put target words in utterance-final position. Even though this finding has often been referred to as a major characteristic of IDS word order, our findings suggest that this phenomenon might not hold cross-linguistically. Specifically, our findings suggest that both Dutch and Mandarin Chinese mothers put more target words in isolation in IDS compared to ADS, but there is no evidence to suggest that they put more target words in utterance-final position in IDS compared to ADS.

The current study focused on the role of speech register and the familiarity of words to children, although many other factors may influence speaking rate, for example, word frequency, word position in an utterance, word type, utterance length, and information status (Seifart et al., 2018). As we collected semi-spontaneous speech data from mother-child interactions, it was not possible to control for all the factors. To our knowledge, only one study has controlled for word position by analyzing read speech. In that study, Ko and Soderstrom (2013) had theater students produce 6 sentences with different registers, focus types, and sentence modes, and they found that IDS was slower than ADS across the entire utterances, though the utterance-final words were lengthened to a larger degree. However, this speech pattern may differ from naturalistic maternal input. Another factor that may confound with the effect of familiarity is word complexity. In our Dutch stimuli, the unfamiliar target words were more complex than the familiar words with respect to the number of phonemes. As such, the word complexity may explain the finding that Dutch unfamiliar words were slower than the familiar words. However, if word complexity and word frequency were the only factors that affected speaking rate, one could have expected a significant main effect of Familiarity only (since the unfamiliar words are more complex and less frequent in ADS as well). In our findings, the significant interaction between Familiarity and Condition (ADS/IDS) at the utterance level clearly shows that mothers especially slow down when introducing unfamiliar words (which are also complex and low-frequent) to their children compared to adults. As unfamiliar words to children are generally more complex than children's familiar words, one way to control for the word frequency or word complexity in future studies is to use pseudowords instead of real words. Also, future analysis may take word complexity into consideration to examine whether word complexity alone could account for the results found in this study.

Previous studies on IDS speaking rate often focused on speech addressed to preverbal children, and the slowing down in IDS is often interpreted as evidence for enhancing acoustic details. However, this interpretation would only hold if IDS is consistently slower than ADS across the entire utterances. In fact, Japanese IDS has been shown to not slow down across entire utterances, and it was the final-lengthening that contributed to the slowdown in IDS (Martin et al., 2016). Our results suggest that speakers of Mandarin Chinese IDS do not show evidence of slowing down at all. As such, it is still debatable whether IDS speaking rate may facilitate phonetic learning. Additionally, even if articulation rate in IDS is consistently slowed down, the phonetic properties of slow speech may not enhance phonetic contrasts. For example, VOT (Voice Onset Time), an acoustic feature that is crucial to distinguish voiceless and voiced consonants, is typically long in voiceless consonants, but short in voiced consonants. If VOT in IDS is longer for voiceless consonants, but shorter for voiced consonants, the contrasts between voiceless and voiced consonants are considered enlarged. However, McMurray, Kovack-Lesh, Goodwin, and McEchron, (2013) found that VOTs were lengthened for both voiced and voiceless consonants in IDS, possibly due to a slow articulation rate. In their study, the CV ratio was the same in both ADS and IDS, suggesting that VOTs were affected by the slower articulation rate in IDS and therefore the changes did not necessarily support phonetic categorization. In sum, it is unclear whether phonetic details are exaggerated in IDS due to slower speaking rate, and whether IDS speaking rate would consequently enhance perceptual development.

Instead of examining the role of IDS in phonetic categorization in the first year of life, our study tried to gauge the facilitating effects of IDS from the perspective of word learning. Specifically, we showed that Dutch mothers use a slower articulation rate to highlight unfamiliar words when addressing children. This suggests that Dutch mothers are sensitive to whether their child knows a word or not, and that they use speaking rate to highlight unfamiliar words, which may potentially facilitate their child's word learning. However, our results can only be interpreted as evidence that mothers make use of articulation rate to differentiate familiar and unfamiliar words in the input. Our results cannot account for whether or not children indeed benefit from this speech pattern. To our knowledge, only one study has studied how varied speaking rate affects 11- to 14-month-old children's word segmentation (Wang, Llanos, & Seidl, 2017). A future direction would be to examine whether children are sensitive to decelerated speech when mapping a novel word to an object in word learning experiments, and whether children from different language backgrounds all benefit from such speech pattern.

2.5 Conclusion

Despite robust evidence supporting the universality of IDS, our results suggest that the speaking rate of IDS in word-learning contexts may not be a common characteristic of Dutch and Mandarin Chinese—two typologically distinct languages—using similar speech elicitation methods. Specifically, Dutch IDS was generally slower than Dutch ADS, while Chinese mothers did not seem to slow down in IDS, even when introducing unfamiliar words to children. Also, Dutch mothers slowed down when introducing unfamiliar words in IDS to a greater extent compared to utterances with familiar words. This targeted speech rate pattern in IDS may facilitate children's word learning. Furthermore, articulation rate kept steady from 18 months to 24 months in both languages, showing no evidence of age-related changes. In sum, IDS does not seem to be always slower than ADS. As such, the nature of IDS in word-learning contexts and the specific cues that may account for the potential facilitative effects of IDS require further examination in a diversity of languages.

Chapter 3

Pitch properties of infant-directed speech specific to word-learning contexts: A cross-linguistic investigation of Mandarin Chinese and Dutch*

Abstract

This study investigates prosodic input, specifically pitch properties, in word-learning contexts in which mothers introduce unfamiliar words to children. We examined Mandarin Chinese and Dutch infant-directed speech (IDS) addressed to 18- and 24-month-old children. Using a semi-spontaneous storybook-telling task, we examined (1) whether mothers made distinctions between unfamiliar and familiar words with pitch in IDS compared to adult-directed speech (ADS); (2) whether IDS prosody changes when mothers address children from 18 to 24 months; and (3) how Mandarin Chinese and Dutch IDS differ in their pitch properties in word-learning contexts in which mothers introduce unfamiliar words to children. Results show that the mean pitch of Mandarin Chinese IDS was already ADS-like when children were 24 months, but Dutch IDS remained exaggerated in pitch at the same age. Crucially, Mandarin Chinese mothers used a higher pitch and a larger pitch range in IDS when introducing unfamiliar words while Dutch mothers used a higher pitch specifically for familiar words. These findings suggest that the prosody of IDS specific to word-learning contexts differs between Dutch and Mandarin Chinese. Based on these findings, whether IDS prosody has linguistic functions and to what extent prosodic cues have potential facilitative effects on word learning requires further examination across languages.

* An adapted version of this chapter is under revision for a journal.

3.1 Introduction

Infant-directed speech (IDS) is an important type of input in early language acquisition (Ramírez-Esparza, García-Sierra, & Kuhl, 2014). Prototypical IDS has exaggerated prosody compared to adult-directed speech (ADS), and is often considered universal across languages and cultures (see reviews in Cristia, 2013; Soderstrom, 2007). IDS prosody has been shown to facilitate word learning; toddlers learn words better from IDS compared to ADS (Graf Estes & Hurley, 2013; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011), and some aspects of IDS have been associated with children's vocabulary size (e.g., Kalashnikova & Burnham, 2018; Porritt, Zinser, Bachorowski, & Kaplan, 2014). However, the language-specificity of IDS is often neglected in the literature (see Wang, Seidl, & Cristia, 2016 for a review). Also, no study to date has specifically investigated the prosody of IDS in word-learning contexts in which mothers introduce unfamiliar words to children. Specifically, it is not clear whether mothers use prosody to highlight novel (unfamiliar) words compared to known (familiar) words when addressing children. Furthermore, as most studies focus on IDS addressed to preverbal children, less is known about the age-related changes of IDS prosody in the second year of life when vocabulary learning is accelerated (Goldfield & Reznick, 1990). To better understand the language-specificity of IDS prosody in word-learning contexts and to demonstrate the age-related changes of IDS prosody in the second year of life, we conducted a cross-linguistic investigation of IDS using similar speech elicitation methods in two different languages. Since we focus on pitch cues relevant to word learning in linguistic input, we chose two languages that differ in their use of pitch at the lexical level, namely Mandarin Chinese, a tonal language, and Dutch, a non-tonal language.

3.1.1 IDS facilitates lexical development

Typically developing children acquire their vocabulary at a fast speed in the first two years of life. They recognize some common words at 6–9 months (Bergelson & Swingley, 2012), start to produce words by the end of their first year (Bloom, 2001), and become proficient word learners at around 18 months old. From about 18 months to 24 months, children's word learning ability gradually improves and their vocabulary size rapidly increases (Bion, Borovsky, & Fernald, 2013; Goldfield & Reznick, 1990). In order to learn words, children need to be familiar with the sounds in their native language(s) and must be able to segment words from continuous speech, recognize familiar words in speech, and associate a novel word label to an object or an action.

Children learn words from language input, however little is known about the quality of prosodic input in word-learning contexts, in which mothers introduce unfamiliar words to their child. IDS is an important type of input in early language acquisition, which has a distinctive prosody compared to ADS. Prototypical IDS is mainly characterized by a higher pitch, a larger pitch range, and a slower speaking rate (Cristia, 2013). These prosodic modifications in IDS have been shown to attract infants' attention, convey positive affect, and facilitate language acquisition (Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002; Soderstrom, Blossom, Foygel, & Morgan, 2008). In order to investigate the role of IDS in lexical development, studies often compare the prosodic characteristics of IDS with those of ADS (see Thorson, 2018 for a review).

Two lines of studies have shown that the prosody and, in particular, the pitch properties of IDS may play a significant role in children's lexical development. The first line of research shows evidence that the pitch properties of IDS correlate with children's vocabulary size, and the second line directly compares children's word learning performance under either ADS or IDS conditions. With regard to the first line of research, only a few studies have investigated the correlations between IDS pitch and language outcomes, and the results are inconsistent. For instance, Porritt et al., (2014) found that English-speaking mothers who had a higher F0 range in their speech had children with larger expressive vocabulary. However, in a recent study, Kalashnikova & Burnham (2018) did not find any correlation between the exaggeration of pitch in Australian English IDS and children's vocabulary size. As the authors suggested, the pitch modifications in IDS may not be related to the facilitative effects of IDS on language acquisition. Taken together, whether pitch properties in IDS are related to vocabulary size is still unclear.

The second line of research suggests that children generally perform better in tasks related to lexical acquisition when they hear prototypical IDS compared to ADS. For example, English- and German-learning children could only segment words in continuous speech from IDS input but not when hearing ADS input (Thiessen, Hill, & Saffran, 2005; Mani & Pätzold, 2016). English-learning infants were able to recognize words that they were familiarized with in IDS even after 24 hours, but not when the words were introduced in ADS (Singh, Nestor, Parikh, & Yull, 2009). When it comes to word-to-object mapping, Ma, Golinkoff, Houston, & Hirsh-Pasek (2011) showed that English-learning 21-month-old children succeeded at word-to-object mapping when listening to the auditory forms of words presented in IDS but not in ADS. Only after 27 months of age could children learn novel words presented in ADS. Similarly, Graf Estes and Hurley (2013) found that 17.5-month-old English-learning children only learned word-object pairings when the words were produced in IDS, but they failed to learn words in the ADS condition. The facilitative effects of IDS on word learning are not restricted to behavioral evidence. Zangl and Mills (2007) found that IDS increased infants' neural activity compared with ADS. Specifically, 6-month-old English-learning children only showed increased neural activity in response to familiar words in IDS compared to ADS, and when children reached 13 months of age they showed increased neural activity for both familiar and unfamiliar words in IDS. It should be noted that the prototypical IDS stimuli in these studies had both exaggerated pitch and a slower speaking rate compared to the ADS stimuli. Thus, while the studies illustrated above invariably suggest that prototypical IDS facilitates children's online word processing, it is not clear whether such facilitative effects can be (solely) attributed to exaggerated pitch. Song et al. (2010) further investigated which acoustic cues in IDS might support word recognition. Their findings suggest that slow speaking rate and vowel hyperarticulation, but not wide pitch range, significantly improved children's online word recognition.

Taken together, prototypical IDS facilitates children's online word learning, but the role of exaggerated pitch in these facilitative effects is not clear. Also, studies on the relationship between IDS pitch properties and lexical development are limited and the results are inconsistent. In order to understand the role of IDS in word learning, it is first necessary to examine how mothers highlight unfamiliar words compared to familiar words in natural IDS. However, so far, little is known about the prosodic input in such word-learning contexts.

3.1.2 Pitch properties of IDS specific to word learning contexts

Based on Hyper and Hypo-speech (H&H) (Lindblom, 1990), Fernald (2000, p. 242) suggests that when interacting with children, adults tend to "[…] modify their speech in ways that serve to maximize predictability for the immature listener […]", which may consequently facilitate children's word recognition. For example, both American English and Japanese IDS have more words produced in isolation and have more repetitions (Fernald & Morikawa, 1993). Also, contextually new information is highlighted by prosodic means in IDS (e.g., Fernald & Mazzie, 1991, to be reviewed later). If IDS is adapted in a way that may facilitate word recognition, it is certainly possible that mothers distinguish unfamiliar words and familiar words with prosody in IDS in support of word learning. To our knowledge, however, no research has directly investigated the prosody of IDS in word-learning contexts by comparing how mothers distinguish unfamiliar words and familiar words with prosody when talking to children. It is important to distinguish between words that are new in a discourse context and words that are unfamiliar to the infants. Young children, as early language learners, encounter unfamiliar words on a regular basis, whereas words directed at adults are usually only "contextually new" within a specific conversational context and rarely novel or unfamiliar. Thus, we use the terms "unfamiliar words" to refer to words that children have not acquired, and "contextually new information" to refer to parts of an utterance that are being introduced to the conversation for the first time but are not necessarily unfamiliar to the addressee. Correspondingly, the term "contextually given information" refers to information that has already been established in the discourse context, while "familiar words" refers to words that the child has already acquired into his or her vocabulary. In ADS, speakers usually highlight contextually new information in discourse by increasing pitch and/or enlarging pitch range, while downplaying given information by reducing these prosodic parameters or by using pronouns in the place of lexical forms when mentioning a word for the second time (Chafe, 1976; Gundel, 1999; Halliday, 1967). In IDS, however, mothers usually repeat the same word several times when talking to children instead of replacing the word with a pronoun (Fernald & Simon, 1984).

Several studies have shown that such prosodic marking of contextually new words is also present in IDS, and its manifestation is different from ADS. Fernald and Mazzie (1991) were the first to examine how F0 is used to highlight contextually new words in English IDS compared to ADS. The target words used in the study were common clothing words (e.g., shorts and socks). To elicit target words, mothers of 14-month-old children were instructed to describe a picture book containing six target items, each introduced successively, to their child and to an adult. They found that mothers typically placed the F0 peak of the utterance on the target words when they introduced contextually new words in IDS; however, the same pattern did not hold true for ADS. Moreover, the second-mention target items (i.e., contextually given words) also showed a greater tendency to occur on F0 peaks in IDS versus ADS. Plus, mothers tended to increase the maximum F0 on the second mentions of a target word compared to the first mentions in IDS. The authors interpreted these results as evidence that prosodic emphasis is placed on both contextually new and contextually given words in IDS—a phenomenon that is not typical in ADS. Even though the authors noted that the familiarity of the target words might vary among the infants, it was not taken into account in their analysis.

Fisher and Tokura (1995) also compared the production of contextually new (first-mention) and contextually given (second-mention) words in English IDS and ADS, but they differed from Fernald and Mazzie (1991) in their elicitation methodology. In this study, mothers of 14-month-old children watched a puppet show which consisted of ten events acted out with ten puppets. The names of the puppets were target words (e.g., tiger, lion, and giraffe). Mothers were asked to describe the events to their child and to an adult. In each event, two puppets were engaged in an action. One puppet (a giraffe) was always in the scene, and the other animal puppet differed across events. The target puppet was always the patient of an action. For example:

> "Your favorite. That's a giraffe. Look he's is petting, and lovin' on the giraffe. Look he's petting' him, pettin' him." (Fisher & Tokura, p. 293)

The prosodic correlates (e.g., F0, position relative to pitch peak, duration, and amplitude) of vowels in the first and second mentions of the target words were analyzed. The results showed that in IDS, vowels in second mentions had a lower pitch, a smaller pitch range, and a shorter duration, indicating that contextually new words were prosodically prominent compared to contextually given words. These results suggest that a given-new contrast does exist prosodically in IDS, however, contrary to Fernald and Mazzie (1991), the authors conclude that the given-new contrast in IDS is similar to the pattern in ADS. As most of the target words were reported as "unknown to the infants" by the mothers (Fisher & Tokura, 1995, p. 292), the familiarity of words was not taken into account as a factor in their analysis.

Since mothers tend to repeat a word several times in IDS, Bortfeld and Morgan (2010) extended Fisher and Tokura (1995) to multiple mentions of target words and examined how mothers of preverbal children (9- and 10-month-olds) mark contextually new and given information across multiple mentions. They used the same methods as in Fisher and Tokura (1995) but conducted their prosodic analyses on entire words instead of vowels. The results showed that when the target words were mentioned for the first time, they received prosodic prominence, while second mentions did not. Specifically, the first mentions showed larger mean F0, higher maximum F0, and longer duration in comparison to second mentions. When measuring more mentions, a significant quartic trend is shown in four acoustic measures: mean F0, maximum F0, F0 range, and duration. These results suggest that mothers alternate between stressed and unstressed realizations across multiple mentions in English IDS. This study, however, did not test ADS in the same task, thus it is not clear whether the same speech pattern would emerge if mothers were involved in the same task in an ADS condition. As in Fisher and Tokura (1995), the familiarity of words was not controlled for in their data analysis.

Even though the results from the studies outlined above are all interpretable as evidence for the highlighting of "new" words in IDS and they all indicate the facilitating effects of IDS on word learning, none of these studies has specifically addressed the prosody of unfamiliar words in comparison with familiar words in IDS.

Despite a lack of understanding about the nature of prosody specific to word-learning contexts in IDS, Grassmann and Tomasello (2007) demonstrated in a word learning study that 24-month-old children learned a novel noun only when it was prosodically accented. In their study, 24-month-old German-speaking children were taught a novel (unfamiliar) noun and a novel (unfamiliar) verb (both of which were phonotactically-legal German pseudo-words) in a sentence, for example "Der Feks miekt," in which either the noun or the verb was accented and marked by a higher pitch, a larger pitch range, and a longer duration. They found that children were able to learn the novel noun when it was both accented and novel but not when it was only accented (but novel) or only novel (but not accented).

To summarize, previous studies have only examined the prosodic marking of contextually new information. The familiarity of words to children has not been taken into account. Consequently, the prosody of IDS in word-learning contexts is not clear from these studies. The current study thus set out to investigate the prosody and specifically the pitch cues of IDS in word-learning contexts. If mothers specifically manipulate pitch in IDS in order to facilitate word learning, they would have an exaggerated pitch (i.e. higher pitch and larger pitch range) when they introduce unfamiliar words as compared to familiar words. Furthermore, it should be noted that previous studies on the prosodic marking of new information were all conducted on English-speaking dyads. It remains unknown whether these results can be generalized to other languages with different prosodic characteristics.

3.1.3 Language-universal and language-specific pitch modifications in IDS

The exaggerated prosody of IDS is found in almost all languages and cultures, with only a few exceptions such as Quiché Mayan (Ingram, 1995; Bernstein Ratner & Pye, 1984). IDS is thus often considered to exist universally across languages and cultures. In most studies on IDS, the speech samples from IDS conditions are natural mother-child interactions or semi-structured play sessions in laboratory settings, while the speech samples from ADS conditions are conversations or interviews with an experimenter. Eliciting speech in such a way ensures the naturalness of speech data, but the content and contexts of speech data in natural mother-child interactions differ to a large extent, making it difficult to directly compare the results between studies on different languages.

Also, cross-linguistic comparisons of IDS are scarce. The few existing cross-linguistic investigations have only examined its generally exaggerated prosody, showing that the differences among IDS in different languages are mainly related to the degree of prosodic exaggeration. For example, even though IDS prosody in all these languages is exaggerated compared to ADS in the same language, the difference in mean pitch between American English ADS and IDS is larger than in British English, French, Italian, German, or Japanese (Fernald et al., 1989). To our knowledge, Grieser and Kuhl (1988) were the first to compare IDS in non-tonal languages (American English and German) and a tonal language (Mandarin Chinese). They found that Mandarin Chinese IDS, as in American English and German, exhibits a higher pitch and larger pitch range compared to Mandarin Chinese ADS. Later, Kitamura et al. (2002) compared the pitch properties (mean pitch, pitch range, and utterance slope-F0) of spontaneous Australian English (a non-tonal language) and Thai (a tonal language) IDS in the first year of life. They found that both Australian English and Thai IDS were more exaggerated than ADS, however, Australian English IDS was generally more exaggerated with respect to pitch properties (mean pitch and pitch range) than Thai IDS. To summarize, cross-linguistic comparisons of IDS in different languages show a universal exaggeration of pitch-related properties compared to ADS, and language-specific aspects seem to be only with respect to the degree of prosodic exaggeration.

However, the prosodic differences between languages may affect IDS in a more complicated way. As mentioned above, previous studies on cross-linguistic comparisons are taken at the general prosodic level without taking a specific context into consideration. In word-learning contexts, different languages may employ different strategies to exaggerate general pitch properties and highlight unfamiliar words while retaining contrastive pitch at the word level. Specifically, IDS in tonal languages and non-tonal languages may show differences in IDS pitch modifications. In non-tonal languages (e.g., English and Dutch), pitch is mainly used for intonational purposes, whereas in tonal languages (e.g., Mandarin Chinese and Thai), pitch is used to distinguish lexical meanings in addition to conveying intonational information. Lexical pitch interacts with the generally exaggerated prosody, which may affect the word and sentence prosody in IDS. This interaction may further impact the pitch in word-learning contexts when unfamiliar words need to be highlighted with pitch on top of the general intonational modifications.

Considering the cross-linguistic differences and the effect of speech contexts on IDS prosody with respect to the different uses of pitch, we set up a word-learning context in which mothers introduced unfamiliar words and familiar words to their child, using similar speech elicitation methods in the two languages: Mandarin Chinese (a tonal language) and Dutch (a non-tonal language).

3.1.4 Age effect

Another factor that affects prosodic modifications in IDS is a child's age. Many studies have investigated the age-related changes of IDS prosody in the first year of life. For example, Stern, Spieker, Barnett, & MacKain (1983) found that the pitch properties in IDS were most exaggerated when children were about 4 months old. Kitamura et al. (2002) investigated the age-related changes in pitch in IDS addressing Australian English learners in their first year of life. They found that the mean F0 increased at 6 months, decreased at 9 months, and increased again at 12 months. However, F0 range did not differ between ADS and IDS in any of the age groups under investigation. In a cross-linguistic comparison of IDS in Korean, Tagalog, and Sri Lankan Tamil, Narayan and McDermott (2016) found that there were no age-related changes from 4 to 16 months. For all the languages, and at all ages under investigation, IDS had a higher pitch and a larger pitch range than ADS. A longitudinal study compared Taiwanese-Mandarin-speaking mothers' speech to preverbal children and to five-year-olds. The degree of pitch exaggeration (measured on vowels) was larger with preverbal children compared to with five-year-old children (Liu et al., 2009).

Most of these studies suggest that pitch-related properties of IDS tend to become less exaggerated as children grow older, though conflicting results exist. Also, most studies focused on the first year of life, thus less is known about how IDS changes beyond the first year. From about 16–18 months to 24 months, both children's receptive and expressive vocabularies start to increase rapidly. This period is known as the "vocabulary spurt" period (Goldfield & Reznick, 1990). Also, during this same age period children's "fast mapping" ability—the ability to map a novel label and a novel object based on minimal exposure—gradually improves. In particular, eighteen-month-old children do not reliably map a novel label to a novel object, but 24-month-old children can reliably associate a novel label to a novel object (Bion, Borovsky, & Fernald, 2013). The current study, therefore, specifically targeted this age range and asked whether Mandarin Chinese and Dutch IDS change from 18 to 24 months.

3.1.5 The current study

Taken together, most studies on IDS to date focus on its general prosody. In particular, no research has addressed whether mothers use pitch to highlight unfamiliar words compared to familiar words in IDS. Plus, there are relatively few cross-linguistic comparisons between IDS in languages with and without lexical tones, and age-related changes of IDS in the second year of life are less understood. Given the potential cross-linguistic differences and age-related changes in the use of pitch in IDS in word-learning contexts, the current study set out to investigate the following research questions: (1) Do mothers make distinctions between unfamiliar and familiar words with pitch in IDS compared to ADS? Specifically, do mothers use an exaggerated pitch (higher pitch and/or larger pitch range) when they introduce words that are unfamiliar to children compared to familiar words in IDS? Since exaggerated pitch attracts infants' attention (e.g., Fernald & Simon, 1984; Masataka, 1992) and children associate novel words with novel objects only when the novel word is marked by a higher pitch, larger pitch range, and longer duration (Grassmann & Tomasello, 2007), we expect that mothers would have a comparatively higher mean pitch and/or a larger pitch range when they introduce unfamiliar words than for familiar words in IDS in order to facilitate word learning. (2) Do pitch properties of IDS and IDS specific to word-learning contexts change when mothers address children from 18 to 24 months? As the prosodic exaggeration usually decreases as children get older, we predict that the global IDS prosody addressing 18-month-old children is more exaggerated than IDS addressing 24-month-old children. Regarding the pitch properties of IDS specific to word-learning contexts, we have two predictions. First, they may become less exaggerated compared to ADS from 18 to 24 months of age in consistency with global pitch modifications. Alternatively, they may remain the same between 18 and 24 months while the global pitch properties become less exaggerated. (3) How do Dutch and Mandarin Chinese IDS differ in their use of pitch cues in word-learning contexts? To answer this research question, we will compare the pitch properties of IDS specific to word-learning contexts in Dutch and Mandarin Chinese.

To address the three research questions, we conducted two experiments using similar materials and procedures in both Experiment 1 (Mandarin Chinese) and Experiment 2 (Dutch).[11] This study is part of a larger study on cross-linguistic comparisons of IDS prosody specific to word-learning contexts (see also Han, de Jong, & Kager, 2018b; Chapter 2 of this dissertation). We adopted a cross-sectional design in the Mandarin Chinese experiment and used a longitudinal design in the Dutch experiment.[12] In both experiments, we used a semi-spontaneous storybook-telling task to elicit both ADS and IDS. The book contains words both familiar and unfamiliar to children. We measured pitch (mean F0 and F0 range) at word and utterance level in the speech data.

3.2 Experiment 1: Mandarin Chinese

3.2.1 Participants

Twenty-one Mandarin-Chinese-speaking mothers of 18-month-old children (mean age = 18;15, age range = 17;21–18;27; girls $N = 9$; mean age of mothers = 30 years,

[11] These two experiments are the same as in Chapter 2, but the order of presentation in this chapter is reversed.

[12] The difference in design was mainly due to the practical situation in which we recruited our participants in China. The participants were mostly recruited from early education programs in kindergartens where they did not enroll for longer than a semester (6 months).

age range = 25–39 years) and nineteen mothers of 24-month-old children (mean age = 24;13, age range = 23;27–24;30; girls *N* = 10; mean age of mothers = 31 years, age range = 32–36 years) participated in the study. All mothers had higher education. The Mandarin Chinese dyads were recruited from kindergartens in Yichang, China. All the participant mothers spoke Mandarin Chinese (the official language in China) proficiently.[13] All children were typically developing.

3.2.2 Materials

A picture book was designed to elicit a set of seven target words, with five unfamiliar words and two familiar words (**Table 3.1**). For each page, one target word was shown on the left side, and an illustration including a depiction of the target word was shown on the right side (see **Appendix A** for example pages of the picture book). Aside from the target words, no other script was provided. An additional six pages of pictures were used as fillers throughout the book to make the story coherent. The target words were all disyllabic nouns. As we wanted to use similar materials for both the Mandarin Chinese and Dutch experiments, we selected familiar words that were listed in both the Mandarin Chinese (M-CDI, Tardif, Fletcher, Liang, and Kaciroti, 2009) version and the Dutch version (N-CDI, Zink & Lejaegere, 2002) of MacArthur-Bates Communicative Development Inventories (CDI, Fenson, Bates, Dale, Marchman, & Reznick, 2007). In contrast, the unfamiliar words were not listed in either M-CDI or N-CDI. Also, the familiar words were more frequent than the unfamiliar words in each language.[14] Selecting target words in such a way was to ensure that the default

[13] All the participant mothers spoke Mandarin Chinese and a dialect (Southwest Mandarin). The participant children heard this dialect in their language community, but were exposed to Mandarin Chinese at home, at kindergarten, and in the national media. This type of bilingual language background is common for most people in China (Li & Lee, 2006). We set these criteria when recruiting participants: (1) the mothers should speak Mandarin Chinese with good proficiency; (2) the mothers should mostly speak Mandarin Chinese to their children at home; and (3) the children should be learning Mandarin Chinese as one of their first languages.

[14] The ranking (lower rank indicating a higher frequency) of Mandarin Chinese word frequency based on Cai and Brysbaert (2010) is: *yé ye* (“grandpa”) (1662), *píng guǒ* (“apple”) (2939), *mí lù* (“moose”) (17914), *hé lí* (“beaver”) (55578), *hé tao* (“walnut”) (12883), *chéng bǎo* (“castle”) (3149), and *nán guā* (“pumpkin”) (5744). The ranking of Dutch word frequency according to Keuleers, Brysbaert, and New (2010) is: *opa* (“grandpa”) (1211), *appel* (“apple”) (4666), *eland* (“moose”) (12385), *bever* (“beaver”) (11515), *walnoot* (“walnut”) (28953), *kasteel* (“castle”) (2185), *pompoen* (“pumpkin”) (12830), *bamboe* (“bamboo”) (30072), *wezel* (“weasel”) (14576), *emoe* (“emu”) (76161), *kapel* (“chapel”) (8604), *jasmijn* (“jasmine”) (26190). Note that word frequency is only provided to show that unfamiliar words usually have a lower word frequency. Ranking is not comparable between languages. We used the mothers’ reports as an indication for Familiarity in analyses.

familiarity of the words applied to most of the participants. However, due to individual differences in vocabulary knowledge, the actual familiarity of the target words might vary among children. Thus, after reading the picture book in both ADS and IDS conditions, mothers filled out a word checklist after the experiment to determine whether their child had already understood the target words before the experiment. This information was coded as Familiarity (Familiar/Unfamiliar) and used in data analyses.

Table 3.1. Target words in Experiment 1 and Experiment 2.

Default Familiarity	**Dutch 18 months**	**Chinese 18 and 24 months (*Pinyin*)**	**English translation**	**Dutch 24 months**	**English translation**
Familiar	opa	yé ye	"grandpa"	opa	"grandpa"
Familiar	appel	píng guǒ	"apple"	appel	"apple"
Unfamiliar	eland	mí lù	"moose"	emoe	"emu"
Unfamiliar	bever	hé lí	"beaver"	wezel	"weasel"
Unfamiliar	walnoot	hé tao	"walnut"	bamboe	"bamboo"
Unfamiliar	kasteel	chéng bǎo	"castle"	kapel	"chapel"
Unfamiliar	pompoen	nán guā	"pumpkin"	jasmijn	"jasmine"

3.2.3 Procedure

All participants were tested in a quiet room. Before the experiment, mothers were given a few minutes to familiarize themselves with the book. Each experiment consisted of two conditions: an IDS condition and an ADS condition. In the IDS condition, the child sat on his or her mother's lap, and the mother was instructed to tell the story to her child the way she usually would at home. The mothers were specifically told that they could use any sentences; the only requirement was to include the words given on each page. In the ADS condition, the mothers were instructed to tell the story to the experimenter (female, a native speaker of Mandarin Chinese), and to take into account the fact that she was a college student. The order of the two conditions was counterbalanced across participants. A ZOOM H1 recorder (with 16-bit resolution and a sampling rate of 44.1 kHz) was used to make audio recordings. Each experimental session took about 15–20 minutes. All participants received a book as a gift after the experiment.

3.3 Experiment 2: Dutch

3.3.1 Participants

Thirty[15] Dutch-speaking mother-child dyads participated when children were 18 months old (mean age of children = 18;14, age range = 18;00–18;29; girls *N* = 17; mean age of mothers = 35 years, age range = 29–44 years). The same participants visited the lab again when the children were 24 months old (mean age of children = 24;18, age range = 24;00–26;30). The Dutch mother-child dyads were recruited from the Utrecht Baby Lab database and were all Dutch native speakers living in the Utrecht area in the Netherlands. As in the Experiment 1, all mothers had higher education and all children were typically developing.

3.3.2 Materials

For the Dutch 18-month-old and 24-month-old children, two picture books were designed to elicit two sets of seven target words, with five unfamiliar words and two familiar words in each set (**Table 3.1**). The book and the target words for Dutch 18-month-old children was the same with the Mandarin Chinese version. To ensure that children had not learned the words at 24 months, the five unfamiliar words in the 24-month-old version were replaced with new unfamiliar words, while keeping the book structure similar for both age groups.

3.3.3 Procedure

All participants were tested in a quiet room in the Utrecht Baby Lab. Each mother-child dyad came to the lab twice, once when the child was 18 months and once when the child was 24 months. The procedure was similar to Experiment 1, however, the experimenter was a native Dutch speaker (female).

3.4 Data analysis

A trained Mandarin Chinese native speaker (the author) and a Dutch native speaker annotated and extracted the target words and target utterances (utterances containing the target words) from the recordings using Praat (Boersma & Weenink, 2017). An utterance boundary was defined in accordance with Martin et al. (2016) as "any pause longer than 200ms which is preceded by an intonational phrase boundary (pauses not accompanied by an IP boundary were considered utterance internal)" (Martin et al., 2016, p. 54). We followed Bortfeld and Morgan (2010) and extracted a minimum F0, maximum F0, and mean F0 (in Hz) of the target words. We also extracted these values

[15] Compared to the Dutch experiment reported in Chapter 2 (*N* = 32), two participants were excluded due to background noise in the recordings.

from the utterances containing target words (i.e. target utterances). The F0 range was calculated as Maximum F0 – Minimum F0. Following Kitamura et al. (2002), the F0 range was transformed to Semitones (st) using the formula: Semitones = $12*\log_2$(maximum F0/minimum F0). The values were extracted automatically using a Praat script and checked manually for doubling and halving errors.

In total, 1375 Chinese utterances and 1434 Dutch utterances were elicited, among which were 541 familiar utterances in Dutch (ADS: 226) and 857 familiar utterances in Chinese (ADS: 335).

To examine whether mothers heightened pitch and/or enlarged pitch range specifically for unfamiliar words in IDS, we used linear mixed-effects models for all analyses. In the models, we included fixed factors of Age (18 months/24 months), Condition (ADS/IDS), and Familiarity (Familiar/Unfamiliar) with Participant as a random factor. The analyses were performed for each language on both word and utterance levels. In Mandarin Chinese, due to the cross-sectional design, we included Condition and Familiarity but not Age as random slopes. In Dutch, we allowed for random slopes for Age, Condition, and Familiarity (Barr, Levy, Scheepers, & Tily, 2013). The dependent variables were word mean F0 (Hz), word F0 range (Semitone (st)), utterance mean F0 (Hz), and utterance F0 range (Semitone (st)). We used the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in the R environment (R Core Team, 2018). For each dependent variable, we took the backward elimination approach, starting with a model that included all fixed effects plus the random factor, and all interactions between them (the most complex model)[16] (Bates, Kliegl, Vasishth, & Baayen, 2015). Then, we used the "step" function in the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) to reduce the models by eliminating non-significant factors or interactions. When the models with maximal random effects failed to converge, we excluded Age from the random slopes. The means and standard deviations of each dependent variable are presented in **Table 3.2**.

[16] An example of the R codes is: lmer(meanF0 ~ Age * Condition * Familiarity + (1 + Age + Condition + Familiarity | Participant))

Table 3.2. Mean word and utterance mean F0 (Hz) and F0 range (st) in Mandarin Chinese and Dutch, standard deviations in parentheses.

Language	Familiarity	Condition	Word Mean F0	Word F0 range	Utterance Mean F0	Utterance F0 range
Mandarin Chinese 18 months	Familiar	ADS	230.54 (51.78)	8.29 (4.62)	238.07 (39.10)	13.17 (5.04)
		IDS	249.61 (55.92)	9.45 (5.82)	256.46 (45.82)	13.94 (5.52)
	Unfamiliar	ADS	238.98 (47.48)	9.14 (5.05)	236.42 (40.31)	13.23 (5.27)
		IDS	272.91 (53.99)	10.00 (5.53)	270.18 (48.57)	13.73 (5.81)
Mandarin Chinese 24 months	Familiar	ADS	237.32 (49.97)	9.62 (5.60)	247.20 (43.39)	14.87 (5.53)
		IDS	247.60 (55.63)	8.14 (5.03)	256.14 (51.88)	13.10 (5.85)
	Unfamiliar	ADS	245.75 (53.52)	8.62 (4.89)	254.16 (50.51)	13.56 (5.64)
		IDS	252.15 (60.23)	9.54 (5.62)	256.83 (47.72)	13.67 (6.08)
Dutch 18 months	Familiar	ADS	250.49 (66.51)	9.03 (6.56)	226.65 (36.33)	16.13 (5.36)
		IDS	271.38 (88.19)	9.29 (5.45)	273.70 (62.15)	15.38 (5.75)
	Unfamiliar	ADS	242.08 (69.11)	9.25 (5.88)	229.45 (42.87)	15.03 (5.74)
		IDS	257.14 (69.62)	9.92 (5.87)	255.49 (58.65)	15.40 (6.38)
Dutch 24 months	Familiar	ADS	224.86 (62.48)	8.45 (5.33)	225.95 (47.78)	15.29 (5.15)
		IDS	269.61 (83.22)	9.33 (5.77)	259.69 (60.61)	15.61 (5.62)
	Unfamiliar	ADS	219.56 (50.44)	7.65 (5.02)	224.36 (39.96)	14.44 (5.24)
		IDS	244.94 (55.52)	9.72 (5.77)	240.25 (44.39)	14.20 (6.04)

3.5 Results

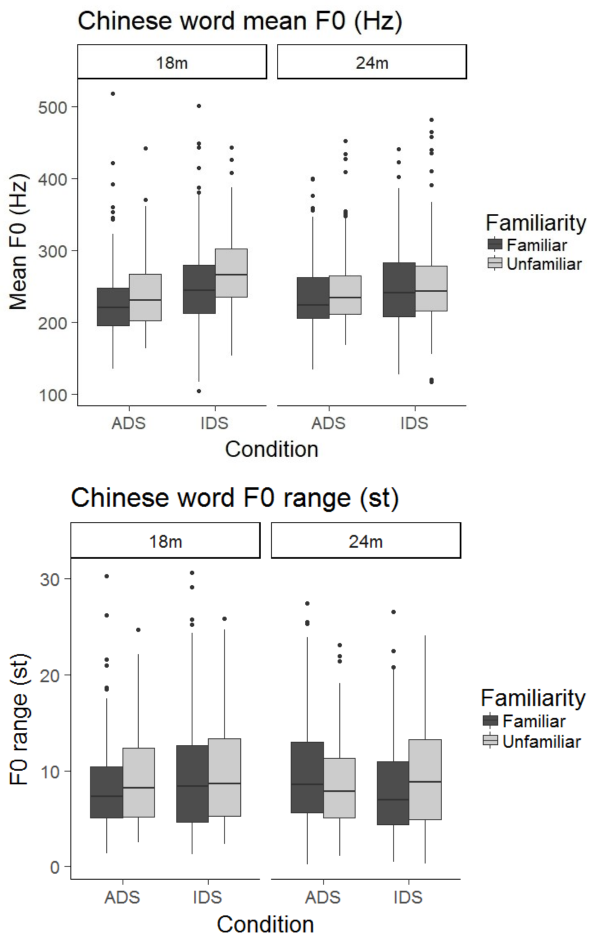
3.5.1 Experiment 1: Mandarin Chinese

We checked whether there was an effect of testing order (ADS-IDS/IDS-ADS) for each dependent measure and no significant differences were found between the two testing orders for any of the dependent measures. Regarding the research questions, we first examined whether unfamiliar words specifically had a higher mean F0 and a larger F0 range than familiar words in IDS as compared to ADS. **Figure 3.1** and **Figure 3.2** show the box plots of mean F0 and F0 range at word and utterance level for Mandarin Chinese.

The results showed that there was a main effect of Condition (ADS/IDS) and a main effect of Familiarity (Familiar/Unfamiliar) on word mean F0 (**Table 3.3a**), but there was no significant interaction between Condition and Familiarity. These results suggest that the target words have a higher mean F0 in IDS than in ADS regardless of Familiarity and the unfamiliar words have a higher mean F0 compared to familiar words regardless of Condition.

As for the dependent variable word F0 range, there was a significant three-way interaction of Condition, Age, and Familiarity in the final model (β = 2.908, *SE* = 1.188, *t* = 2.447, *p* = 0.015). Thus, we split the data by Age (18 months/24 months). The results for the 18 months (**Table 3.3b**) showed that there was a significant main effect of Condition (*p* = 0.008), but neither Familiarity nor the interaction between Condition and Familiarity was in the final model, suggesting that mothers expand pitch range for both familiar and unfamiliar words in IDS when children were 18 months. As for the 24-month-old group (**Table 3.3c**), there was a significant interaction of Condition and Familiarity (*p* = 0.017), but there were no significant effects of either Condition or Familiarity, indicating that Mandarin Chinese mothers specifically expand word F0 range for unfamiliar words in IDS when addressing 24-month-old children.

Figure 3.1. Box plots of word mean F0 (upper panel) and word F0 range (lower panel) for ADS and IDS in Mandarin Chinese.

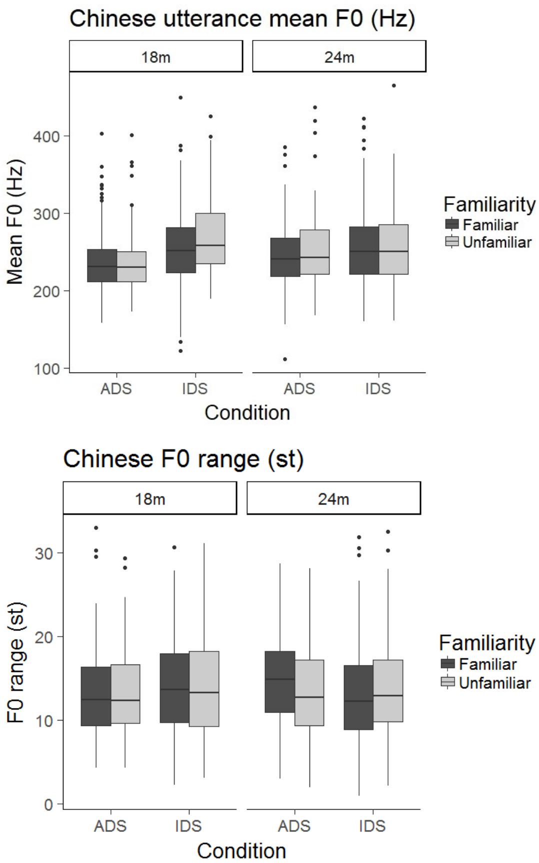


Results at the utterance level showed a significant interaction of Age and Condition on utterance mean F0 (β = -18.766, *SE* = 7.862, *t* = -2.387, *p* = 0.022) and utterance F0 range (β = -1.641, *SE* = 0.605, *t* = -2.712, *p* = 0.007). Thus, we split data by Age for each measurement.

For utterance mean F0 at 18 months (**Table 3.4a**), the results showed that there was a significant effect of Condition and a significant interaction of Condition and Familiarity (β = 15.670, *SE* = 6.312, *t* = 2.482, *p* = 0.013). These results suggest that utterances in IDS had a higher mean F0 compared to ADS, and that this difference was even more pronounced for utterances containing unfamiliar words. The results for 24-month-old children showed that utterance mean F0 did not differ between ADS and IDS, as Condition was not in the final model (**Table 3.4b**).

Now we turn to the results for utterance F0 range. When splitting the data by Age, the final models for utterance pitch range revealed that there was only a significant main effect of Condition for 18-month-old children in IDS (**Table 3.5a**), suggesting that F0 range was larger in IDS as compared to ADS regardless of Familiarity. For 24-month-old children, the final model revealed that there was a significant main effect of Condition, a significant main effect of Familiarity, as well as a significant interaction of Condition and Familiarity (**Table 3.5b**). The direction and size of the effects indicate that, surprisingly, IDS had a smaller pitch range than ADS, and utterances with unfamiliar words had a smaller pitch range than utterances with familiar words. However, the interaction indicates that in IDS compared to ADS, the F0 range of utterances containing unfamiliar words, had a *larger* F0 range.

Figure 3.2. Box plots of Mandarin Chinese utterance mean F0 (upper panel) and utterance F0 range (lower panel) for ADS and IDS in Mandarin Chinese.



Taken together, the results for Mandarin Chinese show age-related changes in IDS prosody. Mandarin Chinese IDS addressing 18-month-old children had a higher mean pitch compared to ADS, but IDS addressing 24-month-old children was already similar to ADS in pitch height. The results also show that Mandarin Chinese mothers tend to use pitch to highlight unfamiliar words. Specifically, at 18 months, when Mandarin Chinese IDS generally had a higher pitch than ADS, utterances with unfamiliar words were specifically higher than utterances with familiar words in IDS. At 24 months, utterance mean pitch of IDS was already similar to ADS, but Mandarin Chinese mothers specifically had a larger word pitch range for unfamiliar words in IDS. Also, utterances containing unfamiliar words had a larger F0 range than utterances containing familiar words specifically in IDS compared to ADS. These findings suggest that Mandarin Chinese mothers of 18- and 24-month-old children distinguish unfamiliar words from familiar words mainly by exaggerating pitch when introducing unfamiliar words.

Table 3.3. Final models for Mandarin Chinese target word mean F0 and F0 range.

3.3a. Final model for Mandarin Chinese target word mean F0 (Hz)				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	228.874	5.351	42.773	<0.001***
Condition (IDS)	18.235	4.751	3.838	<0.001***
Familiarity(Unfamiliar)	13.197	2.902	4.548	<0.001***
	Random factors			
	Variance	***SD***		
Participant(Intercept)	913.2	30.22		
Condition (IDS)	604.5	24.59		
Residual	2181.0	46.70		

3.3b. Final model for Mandarin Chinese target word F0 range (st) for 18-month-old children				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	8.588	0.437	19.657	<0.001***
Condition (IDS)	1.100	0.412	2.673	0.008**
	Random factors			
	Variance	***SD***		
Participant(Intercept)	1.845	1.358		
Residual	27.281	5.223		

3.3c. Final model for Mandarin Chinese target word F0 range (st) for 24-month-old children				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	9.181	0.669	13.724	<0.001***
Condition (IDS)	-1.019	0.714	-1.427	0.162
Familiarity(Unfamiliar)	-0.831	0.609	-1.364	0.173
Condition (IDS): Familiarity (Unfamiliar)	1.936	0.810	2.389	0.017*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	5.252	2.292		
Condition (IDS)	4.156	2.039		
Residual	23.983	4.897		

Note: Intercept in Table 3.3a represents ADS and Familiar. Intercept in 3.3b represents ADS. Intercept in Table 3.3c represents ADS and Familiar. $* p < 0.05$; $** p < 0.01$; $*** p < 0.001$

Table 3.4. Final models for Mandarin Chinese utterance mean F0 for 18-month-old and 24-month-old children.

Table 3.4a. Final model for Mandarin Chinese utterance mean F0 for 18-month-old children

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	233.446	6.233	37.452	<0.001***
Condition (IDS)	20.190	5.406	3.735	<0.001*
Familiarity(Unfamiliar)	-1.490	4.821	-0.309	0.757
Condition(IDS):Familiarity (Unfamiliar)	15.670	6.312	2.482	0.013*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	650.2	25.50		
Condition (IDS)	341.1	18.47		
Residual	1197.9	34.61		

Table 3.4b. Final model for Mandarin Chinese utterance mean F0 for 24-month-old children

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	249.316	7.321	34.05	<0.001***
Familiarity(Unfamiliar)	7.384	3.115	2.37	0.018*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	1217.2	34.89		
Condition (IDS)	562.1	23.71		
Residual	1404.5	37.48		

Note: Intercept in 3.4a represents ADS and Familiar. Intercept in 3.4b represents Familiar. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 3.5. Final models for Mandarin Chinese utterance F0 range.

3.5a. Final model for Mandarin Chinese utterance F0 range (st) for 18-month-old children				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	13.034	0.518	25.152	<0.001
Condition (IDS)	0.812	0.404	2.011	0.045*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	3.572	1.89		
Residual	26.007	5.10		

3.5b. Final model for Mandarin Chinese utterance F0 range (st) for 24-month-old children				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	14.708	0.608	24.209	<0.001***
Condition (IDS)	-1.683	0.603	-2.789	0.005**
Familiarity(Unfamiliar)	-1.435	0.687	-2.089	0.037*
Condition (IDS): Familiarity (Unfamiliar)	1.886	0.903	2.089	0.037*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	2.883	1.689		
Residual	30.922	5.561		

Note: Intercept in 3.5a represent ADS. Intercept in 3.5b represents ADS and Familiar.
$*p < 0.05$; $**p < 0.01$; $***p < 0.001$

3.5.2 Experiment 2: Dutch

We performed similar analyses for the Dutch data. **Figure 3.3** and **Figure 3.4** show the box plots of mean F0 and F0 range for Dutch. We first examined whether unfamiliar words specifically had a higher mean F0 and/or a larger F0 range than familiar words in IDS as compared to ADS. The final model for Dutch word mean F0 (**Table 3.6a**) showed that there were significant main effects of Age and Condition. There were also significant interactions of Age and Condition as well as Condition and Familiarity. For some reason, apparently, the mothers spoke with a lower word mean F0 in ADS when they came back to the lab when their children were 24 months old. In IDS, however, their word mean F0 at 24 months old was higher compared to ADS. Also, unexpectedly, word mean F0 was specifically lower for unfamiliar words in IDS as compared to ADS.

Regarding word F0 range, there was only a significant main effect of Condition. As there was no significant interaction of Condition and Familiarity nor a significant interaction of Condition and Age for word F0 range, these results suggest that target words (regardless of Familiarity or Age) had a significantly larger F0 range in IDS than in ADS (**Table 3.6b**).

Figure 3.3. Box plots of word mean F0 (upper panel) and word F0 range (lower panel) for ADS and IDS in Dutch.

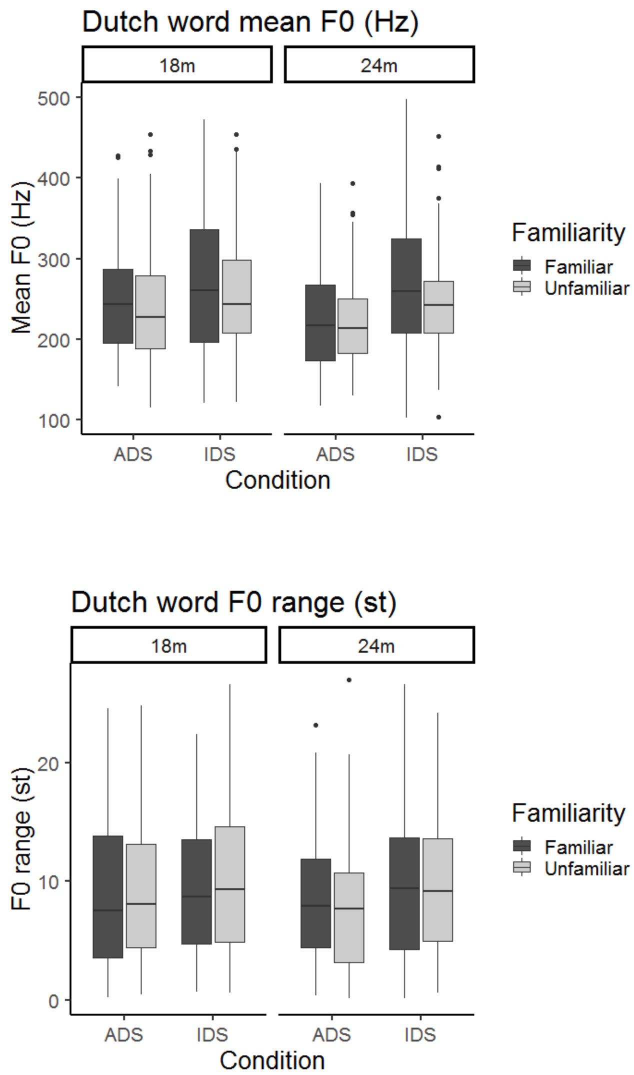


Figure 3.4. Box plots of utterance mean F0 (upper panel) and utterance F0 range (lower panel) for ADS and IDS in Dutch.

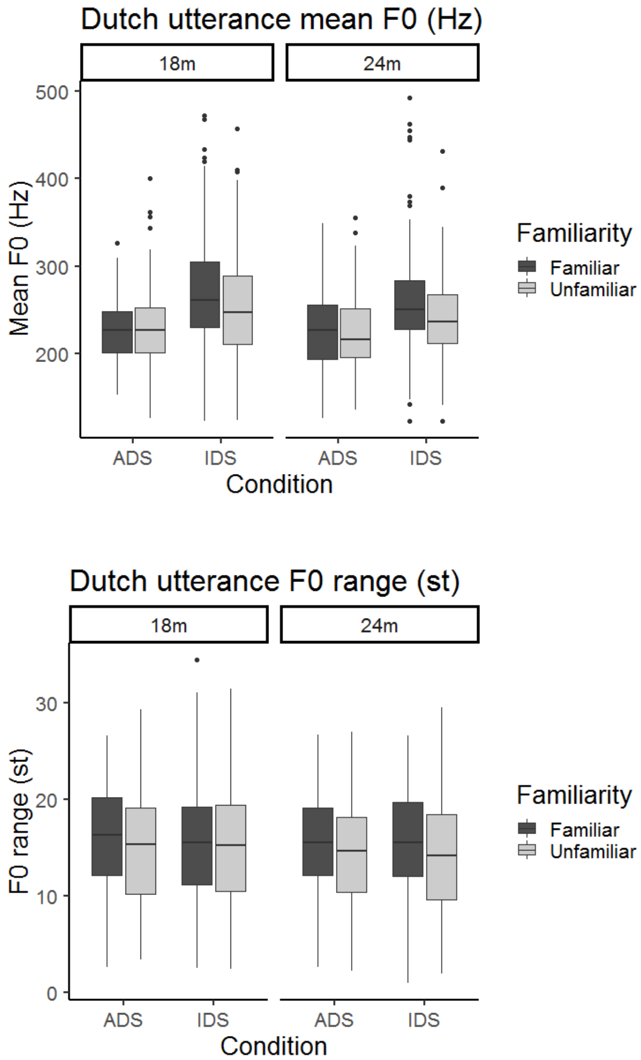


Table 3.6. Final models for Dutch target word mean F0 and F0 range.

3.6a. Final model for Dutch target word mean F0 (Hz)				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	248.188	7.551	32.869	<0.001***
Age (24m)	-23.55	5.27	-4.468	<0.001***
Condition (IDS)	30.479	7.305	4.172	<0.001***
Familiarity(Unfamiliar)	-6.136	6.094	-1.007	0.316
Age (24m) * Condition (IDS)	14.333	6.976	2.055	0.040*
Condition(IDS):Familiarity (Unfamiliar)	-15.813	7.25	-2.181	0.029*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	899.2	29.99		
Condition (IDS)	237.5	15.41		
Familiarity (Unfamiliar)	174.3	13.20		
Residual	4010.0	63.32		
3.6b. Final model for Dutch target word F0 range (st)				
Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	8.6012	0.3675	23.404	<0.001***
Condition (IDS)	1.1216	0.3023	3.711	<0.001***
	Random factors			
	Variance	***SD***		
Participant(Intercept)	2.509	1.584		
Residual	30.208	5.496		

Note: Intercept in 3.6a represents ADS, 18 months, and Familiar. Intercept in 3.6b represents ADS. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$

At the utterance level, the final model showed that there was a significant main effect of Condition, with a significant interaction between Condition and Age as well as Condition and Familiarity (**Table 3.7a**). These results showed age-related changes in IDS: utterance mean F0 was significantly lower when Dutch mothers addressed 24-month-old children compared to 18-month-old children, though utterance mean F0 was higher in IDS compared to ADS at both ages.[17] Also, surprisingly, utterances

[17] When splitting the data by Age, the results showed that for 18 months, there was a significant main effect of Condition ($\beta = 49.10$, SE = 6.61, t = 7.45, $p < 0.001$) and a significant interaction of Condition and Familiarity ($\beta = -20.75$, SE = 7.14, t = -2.90, p = 0.004), but the main effect

containing unfamiliar words specifically had a lower pitch than those containing familiar words in IDS across the two ages.

Table 3.7. Final models for Dutch utterance mean F0 (Hz) and F0 range (st).

Table 3.7a. Final models for Dutch utterance mean F0 (Hz)

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	227.759	6.253	36.426	<0.001***
Age (24m)	-3.111	4.480	-0.695	0.490
Condition (IDS)	49.346	5.635	8.757	<0.001***
Familiarity(Unfamiliar)	-0.127	4.096	-0.031	0.975
Age (24m): Condition (IDS)	-13.777	4.874	-2.826	0.005**
Condition(IDS):Familiarity (Unfamiliar)	-19.573	5.032	-3.890	<0.001***
	Random factors			
	Variance	***SD***		
Participant(Intercept)	796.43	28.22		
Age (24m)	203.19	14.26		
Condition (IDS)	298.89	17.29		
Familiarity (Unfamiliar)	63.27	7.95		
Residual	1902.05	43.61		

Table 3.7b. Final models for Dutch utterance F0 range (st)

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	15.579	0.472	33.029	<0.001***
Familiarity (Unfamiliar)	-0.673	0.308	-2.183	0.029*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	5.109	2.260		
Age (24m)	3.647	1.910		
Residual	27.791	5.272		

*Note: Intercept in Table 3.7a represents ADS, 18 months, and Familiar. Intercepts in 3.7b represents Familiar. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$*

The final model for Dutch utterance F0 range (**Table 3.7b**) showed that there

of Familiarity was not significant (β = -0.269, SE = 5.39, t = -0.05, p = 0.96). Similar results were obtained for the 24-month-old group: There was a significant main effect of Condition (β = 34.49, SE = 6.93, t = 4.97, p < 0.001) and a significant interaction of Condition and Familiarity (β = -17.89, SE = 7.21, t = -2.48, p = 0.013), but the main effect of Familiarity was not significant (β = -0.302, SE = 5.53, t = -0.06, p = 0.96).

was only a main effect of Familiarity, suggesting that utterances containing unfamiliar words had a smaller F0 range compared to utterances with familiar words, regardless of Age or Condition. There were no other significant main effects or interactions.

In sum, our Dutch results show that contrary to our expectations, Dutch mothers had a lower mean F0 specifically for unfamiliar words and utterances containing unfamiliar words in IDS compared to ADS for both age groups. Dutch IDS also enlarged F0 range at the word level compared to ADS, but there were no significant differences in F0 range between ADS and IDS at the utterance level for both age groups. The results also showed age-related changes of mean F0 in IDS: both words and utterances in IDS addressing 24-month-old children had a lower mean F0 compared to IDS addressing 18-month-old children, yet IDS still had a higher mean F0 than ADS for both ages.

3.6 Discussion and conclusions

Despite the robust evidence supporting the universality of IDS, our results suggest that the prosodic input, and in particular the pitch of IDS in word-learning contexts, differs between Mandarin Chinese (a tonal language) and Dutch (a non-tonal language). We conducted two experiments on Mandarin Chinese and Dutch dyads using similar speech elicitation methods. In this design, the content and context were matched between languages as well as between conditions, and we kept the speech data as natural as possible by eliciting semi-spontaneous speech instead of scripted read speech. As the two languages differ in their use of lexical pitch, we focused on the pitch cues in word-learning contexts.

First, we asked whether Mandarin Chinese and Dutch mothers use pitch (e.g., a higher pitch or a larger pitch range) to highlight words that are unfamiliar to children compared to familiar words in IDS. The Mandarin Chinese results confirmed our expectations: when addressing 18-month-old children, utterance mean pitch increased specifically for unfamiliar words in IDS but not for familiar words. At 24 months, both word and utterance pitch range in IDS were exaggerated when mothers introduce unfamiliar words compared to familiar words. However, the Dutch results showed the opposite: Dutch mothers' word and utterance mean pitch raised specifically for familiar words instead of unfamiliar words for both age groups under investigation.

Second, we asked whether IDS prosody changes from 18 to 24 months of age in Mandarin Chinese and Dutch. Our results indicate age-related changes in both languages. Specifically, Mandarin Chinese IDS addressing 18-month-old children had a higher pitch and a larger pitch range than ADS, but IDS addressing 24-month-olds was similar to ADS with respect to mean pitch. Dutch IDS addressing 18- and 24-month-old children both had a higher pitch than ADS, while the pitch range did not differ between IDS and ADS. The results on Dutch pitch range are in accordance with previous findings for Australian and Thai which showed that pitch range did not differ between ADS and IDS (Kitamura et al., 2002). The degree of pitch modifications, indicated by a relatively lower pitch level, was smaller in Dutch IDS addressing 24-month-old children compared to Dutch IDS addressing 18-month-old children. The general trend is that IDS becomes less exaggerated and more ADS-like from 18 months to 24 months in both languages. Previous studies on Taiwanese Mandarin and Dutch IDS focused on the first year of life and their findings showed that IDS had a higher pitch and larger pitch range in both languages (Liu et al., 2009; Van de Weijer, 1999). Our results extend children's age to 24 months by showing that pitch in both Mandarin Chinese and Dutch IDS remains exaggerated compared to ADS until at least 18 months old, well beyond the first year of life. Previous studies that have examined the age-related changes in IDS have generated mixed results for different languages and different age groups under investigation (e.g., Kitamura et al., 2002 on Australian English and Thai; Narayan & McDermott, 2016 on Korean, Tagalog, and Sri Lankan Tamil). Our results contribute to the literature by showing age-related changes in Mandarin Chinese and Dutch IDS from 18 and 24 months.

Third, we asked how Dutch and Mandarin Chinese IDS differ in their use of pitch in word-learning context. Previous studies have shown that IDS has an exaggerated prosody compared to ADS across languages, only the degree of prosodic exaggeration in IDS differs among languages. For example, American English IDS was more exaggerated than British English, French, Italian, German, Japanese IDS (Fernald et al., 1989); Thai IDS was less exaggerated than Australian English IDS (Kitamura et al., 2002). However, as illustrated above, our findings indicate that Mandarin Chinese mothers exaggerate pitch when they introduce unfamiliar words, whereas Dutch mothers exaggerate pitch when they introduce familiar words. These findings suggest that the cross-linguistic differences in IDS are not restricted to the degree of prosodic modifications. In fact, Mandarin Chinese and Dutch mothers exhibit different prosody when introducing unfamiliar words and familiar words to children. Previous studies suggest that mothers are aware of children's vocabulary knowledge at an item level (Styles & Plunkett, 2009; Fenson et al., 2007). Our findings further implicate that both Dutch and Mandarin Chinese mothers keep track of children's vocabulary knowledge in mother-child interactions and adapt their use of pitch accordingly, as shown by significant interactions of Condition and Familiarity. However, the effect of Familiarity on IDS prosody differs in the two languages. As such, pitch functions differently in Mandarin Chinese and Dutch, and languages employ different means in highlighting unfamiliar words in IDS, which may in turn influence children's strategies for word learning in meaningful ways.

The first question that arises given these results is why Mandarin Chinese and Dutch mothers exhibit completely different prosodic modifications regarding the familiarity of words. First, exaggerated pitch draws children's attention (Fernald & Simon, 1984). Also, children are sensitive to the mapping of prosodically-highlighted words and novel objects (Grassmann & Tomasello, 2007). Thus, we interpret the Mandarin Chinese results as evidence for the potential facilitating effects of IDS on word learning. Pitch cues such as higher pitch and larger pitch range in IDS do not only have linguistic functions, but also serve to signify positive affect (Singh, Morgan, & Best, 2002; Trainor, Austin, & Desjardins, 2000). As such, the Dutch results, which showed higher pitch for familiar words, may be attributed to positive affect. In a longitudinal investigation of Dutch IDS addressed to 11- to 15-month-old children, Benders (2013) found that the acoustic properties of vowels in Dutch IDS convey positive affect but do not enhance vowel contrasts, which could consequently facilitate infants' phonetic categorization. The target words in their study included words such as *fiets* ("bike"), *boek* ("book") and *schaap* ("sheep"). Even though the familiarity of these words for each child was unknown, these words were mostly listed in N-CDI (Zink & Lejaegere, 2002), thus they are likely to be familiar to children. It is possible that Dutch mothers show more positive affect when they mention words that are familiar to their child compared to unfamiliar words, e.g. because placing positive affect on unfamiliar words might not be meaningful. In contrast, they might lower pitch for unfamiliar words to show a relatively neutral emotion. However, little is known about whether showing positive affect on familiar words may help or inhibit language learning. Future research may further investigate the emotional affect in word-learning contexts and the possible effects on word learning.

We have shown that Dutch mothers did not seem to exaggerate pitch to highlight unfamiliar words, however, this does not necessarily mean that Dutch mothers do not highlight unfamiliar words at all. Han, de Jong, and Kager (2018b; Chapter 2 of this dissertation) found that Dutch mothers slowed down their utterances when introducing unfamiliar words compared to utterances containing familiar words in IDS. Combining these results, Mandarin Chinese and Dutch IDS employ different prosodic cues to highlight unfamiliar words. Mandarin Chinese IDS mainly uses exaggerated pitch, while Dutch IDS prefers temporal cues (i.e. articulation rate). However, these results only demonstrate how mothers use prosody to make distinctions between unfamiliar and familiar words during mother-child interactions. Future studies should examine whether such speech patterns in Mandarin Chinese and Dutch IDS indeed facilitate word learning in Mandarin Chinese and Dutch children.

The differences in the use of pitch cues in Mandarin Chinese and Dutch may also be attributed to typological differences between these two languages. Mandarin Chinese, as a tonal language, uses pitch to distinguish lexical meanings. As a result, the pitch range of words is crucial to word meanings, so mothers specifically enlarged pitch range when introducing unfamiliar words. They specifically did so at 24 months old, when children are learning words efficiently. Dutch, a stress language, may resort to temporal cues to highlight unfamiliar words.

Existing studies on Dutch or Mandarin Chinese IDS have focused on general prosodic exaggeration and vowel hyperarticulation. At the intonational level, both Dutch and Mandarin Chinese, as in many other languages, have a higher pitch and a larger pitch range compared to ADS when addressing preverbal children (Van de Weijer, 1999; Grieser & Kuhl, 1988; Liu, Tsao, & Kuhl, 2009). Vowels in Mandarin Chinese IDS are hyperarticulated (Liu, Tsao, & Kuhl, 2009; Tang, Xu Rattanasone, Yuen, & Demuth, 2017), but vowels in Dutch IDS show *hypoarticulation* (Benders, 2013). In addition, lexical tones in Mandarin Chinese IDS are hyperarticulated (Han, de Jong, & Kager, 2018a; Tang et al., 2017). However, none of these studies have examined the word and sentence prosody in word-learning contexts. Our study contributes to the understanding of Dutch and Mandarin Chinese IDS in word-learning contexts as well as the age-related changes from 18 months to 24 months.

To conclude, despite robust evidence supporting the universality of IDS, our results suggest that the pitch properties in IDS specific to word-learning contexts differ between Mandarin Chinese, a tonal language, and Dutch, a non-tonal language. Specifically, speakers of Mandarin Chinese IDS enlarge pitch range when they introduce unfamiliar words, but Dutch IDS speakers heighten pitch specifically when introducing familiar words. It is possible that the pitch cues in Mandarin Chinese IDS have more pedagogical functions, while the pitch cues in Dutch IDS convey positive affect and are more entertaining. Furthermore, the developmental changes from 18 months to 24 months differ in these two languages. Both Mandarin Chinese and Dutch IDS are exaggerated compared to ADS in these languages when addressing 18-month-old children. When children reach 24 months, Mandarin Chinese IDS is already similar to ADS, whereas Dutch IDS is still more exaggerated than ADS.

Our study contributes to the understanding of the quality of prosodic input in two distinct languages and cultures. Our findings indicate that the prosodic input in word-learning contexts differs between languages, and consequently, the specific prosodic cues that account for the potential facilitative effects of IDS require further examination in a diversity of languages and cultures.

Chapter 4

Lexical tones in Mandarin Chinese infant-directed speech: Age-related changes in the second year of life*

Abstract

Tonal information is essential to early word learning in tone languages. Although numerous studies have investigated the intonational and segmental properties of infant-directed speech (IDS), only a few studies have explored the properties of lexical tones in IDS. These studies mostly focused on the first year of life; thus little is known about how lexical tones in IDS change as children's vocabulary acquisition accelerates in the second year (Goldfield & Reznick, 1990). The present study examines whether Mandarin Chinese mothers hyperarticulate lexical tones in IDS addressing 18- and 24-month-old children—at which age children are learning words at a rapid speed—vs. adult-directed speech (ADS). Thirty-nine Mandarin Chinese–speaking mothers were tested in a semi-spontaneous storybook-reading task, in which they told the same story to their child (IDS condition) and to an adult (ADS condition). Results for the F0 measurements (minimum F0, maximum F0, and F0 range) of tone in the speech data revealed a continuum of differences among IDS addressing 18-month-olds, IDS addressing 24-month-olds, and ADS. Lexical tones in IDS addressing 18-month-old children had a higher minimum F0, higher maximum F0, and larger pitch range than lexical tones in ADS. Lexical tones in IDS addressing 24-month-old children showed more similarity to ADS tones with respect to pitch height: there were no differences in minimum F0 and maximum F0 between ADS and IDS. However, F0 range was still larger. These results suggest that lexical tones are generally hyperarticulated in Mandarin Chinese IDS addressing 18- and 24- month-old children despite the change in pitch level over time. Mandarin Chinese mothers hyperarticulate lexical tones in IDS when talking to toddlers and potentially facilitate tone acquisition and word learning.

* An adapted version of this chapter has been published as: Han, M., de Jong, N. H., & Kager, R. (2018a). Lexical tones in Mandarin Chinese infant-directed speech: Age-related changes in the second year of life. *Frontiers in Psychology*, *9*, 434.

4.1 Introduction

In tone languages, pitch is employed to differentiate lexical meanings. Consequently, in order to recognize or learn a word, a tone-language-learning infant must develop sensitivity to lexical pitch contours in addition to consonants and vowels; conversely, infants who learn non-tone languages need to pay attention to consonants and vowels but ignore pitch contours at the lexical level. Though a number of studies have looked at infants' discrimination, recognition, and acquisition of tones (see Singh & Fu, 2016 for a review), only a few studies have examined lexical tones in early language input—i.e., infant-directed speech (IDS). The results drawn from these studies are inconsistent; some suggest that tones in IDS are *hypoarticulated*, while others show that they are *hyperarticulated* compared with tones in adult-directed speech (ADS). Moreover, most previous studies have focused on IDS in the first year of life, when perceptual reorganization is taking place (Werker & Tees, 1984); comparatively little is known about how tonal input changes in the second year, when children start to become verbal and gain vocabulary at a rapid speed (Goldfield & Reznick, 1990). As tonal information is crucial to distinguishing word meanings, the current study investigates whether lexical tones in Mandarin Chinese IDS addressing 18-and 24-month-old children are hyperarticulated—and if so, whether the tonal cues change depending on the age of the child.

Infant-directed speech is a speech register caregivers (typically mothers) use when addressing their infants, and as such it is an important type of input in early language acquisition (Cristia, 2013; Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015; Soderstrom, 2007). IDS is known to exhibit exaggerated intonation compared with ADS, including higher pitch, a larger pitch range, and greater pitch variations (Fernald & Simon, 1984; Fernald et al., 1989). These types of prosodic modifications are found in IDS in the majority of world languages, including both non-tone languages, such as English and German (Fernald & Simon, 1984; Fernald et al., 1989; Cristia, 2013), and tone languages, such as Mandarin Chinese, Cantonese, and Thai (Grieser & Kuhl, 1988; Kitamura et al., 2002; Xu Rattanasone et al., 2013). Despite the near-universality of exaggerated intonation, the degree of exaggeration may show cross-linguistic or cross-cultural differences. For instance, American English IDS was found to exaggerate prosody more than British English, Japanese, German, French, and Italian IDS (Fernald et al., 1989). In the IDS of tone languages, lexical tone (pitch at the lexical level) interacts with exaggerated intonation (pitch at the intonational level); as a result, the prosodic modifications expressed in tone-language IDS may differ in meaningful ways from those found in non-tone-language IDS. For instance, Kitamura et al. (2002) found that, although Thai IDS exhibited exaggerated intonation compared with Thai ADS, it was less exaggerated than Australian English IDS.

IDS is often claimed to facilitate language acquisition, although conflicting views have been proposed (Gleitman et al., 1984; Soderstrom, 2007). One line of research has shown that, compared with ADS, IDS attracts infants' attention more effectively. Infants—even newborns—prefer listening to IDS over ADS (Cooper & Aslin, 1990). This listening preference is probably largely attributable to the positive affect of IDS (Singh et al., 2002). Positive affect is a common characteristic of IDS, and one that shares similar prosodic features with exaggerated intonation (Kitamura & Burnham, 1998). When they manipulated affect and speech register in IDS to examine 6-month-old children's listening preference, Singh et al. (2002) found that higher pitch and greater pitch variation alone did not account for infants' preference; positive affect was also required.

The robust evidence that infants prefer listening to IDS, however, does not necessarily indicate that such speech carries a particular linguistic function in terms of language learning. Another line of research has been devoted to identifying the well-specified linguistic information encoded by IDS. A number of studies have explored two questions on this topic: First, are the segmental (mainly vocalic) and suprasegmental (tonal) properties of IDS hyperarticulated compared with those of ADS? It may seem, on first blush, that the exaggerated intonation IDS entails vowel hyperarticulation. However, it is also possible that exaggerated intonation co-occurs with more variable vowels, and thus poses a learning problem for vowel categorization. Similarly, exaggerated intonation need not naturally result in tone hyperarticulation; on the contrary, it may distort tonal cues at the syllabic level. Second, if the segmental and suprasegmental properties of IDS are indeed hyperarticulated, is this hyperarticulation expressed in a way that may support language acquisition? Previous investigations into this possibility have produced mixed results on the segmental level (vowels and consonants) and few results of any kind on the suprasegmental level (lexical tones).

An example of vowel hyperarticulation was identified by Kuhl et al. (1997), who compared the articulation of three point vowels (/i/, /a/, and /u/) between ADS and IDS addressing 2- to 5-month-old infants in American English, Russian, and Swedish. They analyzed the "vowel triangles" for the three vowels in IDS and ADS; a larger vowel triangle indicated that the vowels were more distinctive from each other. The results showed that in all three languages, mothers expanded the vowel triangles in IDS compared with ADS, suggesting that mothers produced more distinctive vowels in IDS. Similar results have been obtained in other languages, including Taiwanese

Mandarin (Liu et al., 2003), French, and Japanese (Dodane & Al-Tamimi, 2007). However, contradictory findings have also been reported. First, vowel hyperarticulation seems to be restricted to point vowels (/i/, /a/, and /u/); when comparing other vowel contrasts such as [i – ɪ] in American English, Cristia and Seidl (2014) did not find these contrasts to be enhanced in IDS. Second, while robust evidence of vowel hyperarticulation exists for multiple languages, other languages seem to show no trace of this phenomenon. For example, vowels in Cantonese IDS toward 3- to 12-month-old infants were not hyperarticulated compared with vowels in Cantonese ADS (Xu Rattanasone et al., 2013). Similarly, a recent study comparing the vowels in natural Japanese IDS addressing 18- to 24-month-old children with the vowels in read Japanese speech found that, although the IDS vowels were more variable, they did not necessarily show more clarity compared with those in ADS (Miyazawa et al., 2017).

The mixed results on vowel hyperarticulation in IDS are only magnified in studies investigating whether IDS supports language acquisition. On the one hand, Song et al. (2010) showed that vowel hyperarticulation in IDS improved word recognition in 19-month-old children. On the other hand, in a perception study on 6- and 7-month-old children, Trainor and Desjardins (2002) found that the exaggerated pitch contours in IDS helped children's discrimination of vowels, whereas high pitch hampered vowel discrimination. In sum, whether or not vowels in IDS are hyperarticulated—and whether such hyperarticulation, if it exists, helps children's language acquisition—is still debatable.

A similar debate may be extended to tone hyperarticulation. Hypothetically, the exaggerated intonation of IDS might affect tonal properties in two possible ways. Specifically, lexical tones in IDS may either be hyperarticulated or alternatively distorted (hypoarticulated) due to the exaggerated prosody. Two types of acoustic evidence may indicate tone hyperarticulation in IDS. First, tones' acoustic cues may be more prominent in IDS as compared with ADS. For example, as fundamental frequency (F0) is the primary cue to tone in Mandarin Chinese (Howie, 1974), tone hyperarticulation can be indicated by a larger F0 range for Tone 2 (mid-rising tone), Tone 3 (low-dipping tone), and Tone 4 (high-falling tone). Tone 1, a high-level tone, may have a higher F0 in IDS than in ADS. Additionally, tone duration, a secondary cue (Blicher et al., 1990), may also be enlarged in IDS for all four tones. Second, enhancement of tonal contrasts is a possible indicator of tone hyperarticulation in IDS. Such enhancement can be measured by comparing the pitch differences between tone pairs in ADS and IDS, or indicated by a larger tone triangle in IDS (e.g., Tang et al., 2017, to review later). To date, only a handful of studies have looked at lexical tones in IDS. Among the few studies that have performed perceptive or acoustic measurements on lexical tones in IDS, conflicting results emerge.

Results from several studies support the distortion prediction. Papoušek and Hwang (1991) found that tone contours in Mandarin Chinese IDS did not correspond to phonologically expected tone contours. In their study, participants were instructed to produce preselected utterances in role-play contexts, imagining the addressee was a child or an adult. The authors speculated that speakers intuitively sacrificed tonal information at the syllabic level in order to accommodate the IDS intonation. Though the study's results shed light on people's intuitive prosodic tuning when talking to children, they do not tell us much about tone production in natural IDS, when mothers and children interact directly. In a later study, Kitamura et al. (2002) collected IDS data from Thai speakers in a more natural setting. Specifically, the researchers recorded the spontaneous speech of mothers interacting with their children naturally at home, every 3 months, from birth until the infants were 12 months old (IDS condition); they also recorded the same participants interacting with adults (ADS condition). They then asked trained Thai phonologists to judge whether the tones in utterance-initial and utterance-final positions remained identifiable. The results showed that tones were slightly less identifiable in Thai IDS than ADS, especially in utterance-final position.

While these studies suggest that tones may be distorted in IDS compared with ADS, there is also evidence that mothers hyperarticulate tones in IDS. Following the methods in Kuhl et al. (1997), Liu et al. (2007) investigated whether vowel hyperarticulation applied to tones in Taiwanese Mandarin. They performed an acoustic analysis on four Taiwanese Mandarin tones in speech directed at 10- to 12-month-old children. Their stimuli consisted of 12 disyllabic words in which the first syllable (target syllable) varied from Tones 1 to 4 and the second syllable remained Tone 1. In the IDS condition, mothers and their infants played together with pictures or objects corresponding to these stimuli; in the ADS condition, the same mothers talked to an experimenter about the children's interests in these target words. Mean F0, F0 range, and duration of vowels of the target syllables were compared between the two conditions. The results showed that Taiwanese Mandarin tones produced in IDS had a raised mean F0, enlarged F0 range and lengthened duration—suggesting that mothers tended to hyperarticulate tones when speaking to their infants.

Two studies further tested tone hyperarticulation in Cantonese IDS with different measurements. Xu Rattanasone et al. (2013) investigated Cantonese tones in the speech of mothers talking to their 3-, 6-, 9-, and 12-month-old children. The stimuli consisted of three of the six tones in the Cantonese tone inventory: tones 55, 25, and

21. The authors adopted a tone triangle measure from Barry and Blamey (2004). For each tone, F0 values were measured at the point of maximal vowel amplitude and at 50% of the maximum amplitude. These two values were plotted for three tones, making a tone triangle. Similar to the vowel triangles in Kuhl et al. (1997), a larger tone triangle indicated more distinctive tonal contrasts. The results showed that tone triangles were larger in IDS than ADS at 3, 6, and 9 months, indicating tone hyperarticulation for these age groups. However, the observed hyperarticulation was reduced for 12-month-olds, indicating that tones in speech to infants are more distinctive until children reached 12 months of age, at which point tones in ADS and IDS become similar. Significantly, the larger tone triangle found for 3-, 6-, and 9-month-olds mainly stemmed from differences between the high-level tone (55) and the low-level tone (21) (Xu, 2008, p. 111); thus, it remains unknown whether these larger tone triangles indicate tone hyperarticulation across the whole tone inventory. In a recent study, Wong and Ng (2017) examined Cantonese tone hyperarticulation in IDS (toward 7- to 12-month-old infants), using both native judgment and acoustic analysis. They found that tones in Cantonese IDS had higher F0 and longer duration than tones in ADS, but such differences did not seem to facilitate adults' perception of tonal contrasts. Using the tone triangle measure in Xu Rattanasone et al. (2013), Tang et al. (2017) examined tone hyperarticulation in Northern Mandarin. Interestingly, they only found tone hyperarticulation (for both tone space and duration) when the target tones were in utterance-final position.

Taken together, these findings indicate that Cantonese tones are hyperarticulated in early IDS compared with ADS, but that the degree of hyperarticulation diminishes by the end of the first year. They also suggest that tone hyperarticulation may be restricted to certain tones or positions (as in Northern Mandarin, where tone hyperarticulation is only present in utterance-final positions). In other words, it has not been conclusively established that lexical tones in IDS are hyperarticulated across the board. To date, studies of IDS have been conducted on different languages, with different data collection methods and different measurements, and have yielded conflicting results. These methodological issues must be taken into consideration before we draw any conclusions about the hyperarticulation of tone in IDS.

Tone languages studied in the existing literature on IDS include Cantonese, Mandarin Chinese, Northern Mandarin, and Taiwanese Mandarin, all of which have different tonal systems and prosodic patterns (e.g., Chen, Wang, & Xu, 2009). It is possible that the interaction of tone and prosodic modifications in IDS may show cross-linguistic differences. In fact, even among variants of the same language, the characteristics of IDS can differ; for example, as noted above, American English IDS tends to be more exaggerated than British English IDS (Fernald et al., 1989)—certainly implying that languages with different tonal systems or different dialects may differ significantly.

Second, speech elicitation methods used in previous studies range from reading tasks to spontaneous speech, and from home settings to laboratory settings. Papoušek and Hwang (1991) used scripted speech, while Liu et al. (2007) selected target words to elicit speech during mother-child interaction (semi-spontaneous), and Kitamura et al. (2002) collected spontaneous speech data in natural interactions at home. Prosody tends to differ in read speech vs. (semi-)spontaneous speech (de Ruiter, 2015). In spontaneous speech (elicited during "natural mother-child interaction"), the speech context varies according to the activity that is taking place—for example, reading books, playing with toys, or changing diapers. Furthermore, in typical experimental settings, the speech contexts for ADS and IDS conditions are rather different from each other. It is not surprising, then, that IDS may be more distinct from ADS in certain contexts, and less distinct in other contexts. Given this degree of variability, it's not clear whether the large differences between ADS and IDS reported in certain previous studies may actually have been due to the very different settings and activities in the two conditions.

Finally, previous studies have employed a wide range of analyses to compare the ADS and IDS conditions. Kitamura et al. (2002) used native judgment, whereas other researchers performed acoustic analyses; among the studies that conducted acoustic analyses, different measurements were used. These methodological differences further complicate the task of determining whether or not lexical tones are hyperarticulated in IDS.

Besides the methodological issues discussed above, the different ages of the children in the various studies may also have contributed to the contradictory results. Studies on vowel and tone hyperarticulation to date have mostly focused on IDS directed at children in the first year of life, and these results have often been interpreted from the perspective of "perceptual reorganization" (Werker & Tees, 1984). There is robust evidence showing that infants undergo perceptual reorganization, during the first 12 months of life, as their perception of phonetic categories shifts from language-universal to language-specific. This shift is reflected in infants' progressively better discrimination of native contrasts and poorer discrimination of non-native contrasts. Such perceptual reorganization develops for consonants, vowels, and lexical tones. Mandarin-learning infants, for instance, show improvement in their discrimination of lexical tones between 6 and 9 months of age, while infants who are learning a non-tone language (e.g., English and Dutch) show a decline in their ability to discriminate tonal contrasts over the same age range (Mattock & Burnham, 2006; Liu & Kager, 2014). Thus, findings on tone hyperarticulation during infancy are usually interpreted as evidence for the facilitating effects of IDS on tone perception: as infants' speech perception becomes progressively tuned to their native (tonal) language, tone hyperarticulation becomes less prominent. Xu (2008, p. 99), for instance, pointed out that her findings—which indicate that tone hyperarticulation declines at 12 months—are consistent with perceptual reorganization research. However, during the same period of perceptual reorganization, children also start to acquire words. Infants start to show recognition of common words as early as 6–9 months (Bergelson & Swingley, 2012), and usually utter their first words around their first birthday. In the second year of life, both receptive and productive vocabulary accelerate at an astonishing speed (Goldfield & Reznick, 1990).

Since tonal information is crucial to word meaning in tone languages, it is important to examine whether tone hyperarticulation persists when children are becoming proficient word-learners in the second year. The general prosodic modifications in IDS are known to change based on the child's stage of language development (Stern et al., 1983; Kitamura et al., 2002). In general, IDS becomes more ADS-like as children grow older. Taking the perspective of word learning, tone hyperarticulation may not stop when children are one year old; on the contrary, it may persist, aiding children's lexical development as they move into the word-learning phase. As most studies to date have focused on the first year of life, little is known about whether tone hyperarticulation remains present in the second year. Consequently, the timeline of age-related changes in tone hyperarticulation is not well-described in the literature.

Two studies have investigated age-related changes in lexical tones in IDS, but both focused on the first year of life, prior to the lexical spurt. Kitamura et al. (2002) showed that lexical tones in Thai were distorted in IDS directed at children up to 9 months old, but that IDS directed at 12-month-old children did not differ significantly from ADS in tone identification. Results from Xu Rattanasone et al. (2013) showed similar age-related changes: Cantonese tones were hyperarticulated in IDS compared to ADS until 12 months of age, at which point this hyperarticulation was reduced. The authors interpreted their results as evidence that mothers modify their speech according to children's stages of language development. As infants tune their tone perception toward their native language in the first year of life (Mattock & Burnham, 2006; Mattock et al., 2008; Yeung, Chen, & Werker, 2013), tone hyperarticulation declines accordingly.

If age-related changes in IDS are explicitly tied to perceptual reorganization, we should expect any differences between ADS and IDS to diminish and disappear altogether as children reach 12 months of age. However, in a longitudinal study, Liu et al. (2009) found that speech directed to 5-year-old children still showed both general prosodic exaggeration and tone hyperarticulation compared with ADS, though it was less exaggerated than IDS directed at preverbal children. But what happens to IDS directed at children between infancy (up to 12 months) and school-age (5 years old)? There is a gap in the existing investigations of tone hyperarticulation during this period. The present study seeks to fill that gap by asking what happens to tones in IDS in the second year of life, when children start to talk and learn vocabulary at a high rate. It remains an open question whether mothers speaking tone languages alter their tones in IDS to facilitate tone acquisition (and, consequently, lexical development) in their children. If tone hyperarticulation is not restricted to supporting perceptual reorganization, we should find evidence for tone hyperarticulation in IDS addressing 18-and 24-month-olds.

The current study set out to investigate tone hyperarticulation in Mandarin Chinese IDS at two points in time, both of which occur during the second year of life (the period of the lexical spurt). Our main research questions are: (1) Are tones in Mandarin IDS addressed to 18- and 24-month-old children generally hyperarticulated compared to tones in ADS? If so, we should expect to observe a larger F0 range for Tone 2, Tone 3, and Tone 4, a higher F0 for Tone 1, and possibly longer duration for tones in IDS vs. ADS, as shown by Liu et al. (2007). In addition to these general measures, we explored whether lexical tonal contrast between Tones 1 and 4 was enhanced in IDS. (2) Does the hyperarticulation of lexical tones in Mandarin Chinese IDS differ when mothers address 18-month-old versus 24-month-old children? We predict that as children's vocabulary size increases significantly from 18 to 24 months, the lexical tonal cues change. Specifically, tonal cues in IDS should be more similar to ADS when children reach 24 months old. To address these questions, we collected speech samples from a story-telling task, where mothers told a story containing target words featuring four Mandarin Chinese tones to their 18- and 24-month-old children (IDS condition), and to an adult control (ADS condition). The experiment is the same experiment as Experiments 2 and Experiment 1 reported in Chapters 2 and 3, respectively.

4.2 Materials and methods

4.2.1 Participants

Thirty-nine Mandarin-Chinese-speaking mother-child dyads participated in this study. The participant sample comprised two age groups: 18-month-olds ($N = 21$; mean age

= 18;15; age range = 17;21 – 18;27; girls *N* = 9) and 24-month-olds (*N* = 18; mean age = 24;15; age range = 23;27– 24;27; girls *N* = 10). All participants were recruited from kindergartens in Yichang, China. All the participant mothers spoke Mandarin Chinese (the official language in China), as well as a dialect (in this case, Southwest Mandarin).[18] The participant children heard this dialect in their language community, but were exposed to Mandarin Chinese at home, at kindergarten, and in the national media. This type of bilingual language background is common for most people in China (Li & Lee, 2006). To obtain a homogeneous group of participants, we set these criteria in our recruiting interview: (1) the mothers should speak Mandarin Chinese with good proficiency; (2) the mothers should mostly speak Mandarin Chinese to their children at home; and (3) the children should be learning Mandarin Chinese as one of their first languages.

4.2.2 Materials

A picture book titled *Xiaotuzi de yitian* ("Bunny's day") (see **Appendix A** for example pages of the picture book) was designed to elicit four target words for 18- and 24-month-old children (see **Table 4.1**). On each page of the book, one word appeared on the left side, and a corresponding picture appeared on the right side. The pages contained no text beyond these target words. An additional six pages were used as fillers and to make the story coherent. The target words were all disyllabic nouns, of which the first syllable was always Tone 2 (a rising tone), and the second syllables varied from Tones 1 to 4. We chose Tone 2 for the first syllable in order to ensure consistent tonal coarticulation effects (i.e., carry-over effects on the following tone) across tokens and registers.

Table 4.1. Overview of stimuli.

Tone of the second syllables	**Tone 1**	**Tone 2**	**Tone 3**	**Tone 4**
Pinyin	nán guā	hé lí	chéng bǎo	mí lù
IPA	[nan2 kuɑ1]	[xɤ2 li2]	[tʂəŋ2 pɑu3]	[mi2 lu4]
Translation	"pumpkin"	"beaver"	"castle"	"moose"

4.2.3 Procedure

Participants were tested in a quiet room. Before the experiment, mothers were given a few minutes to get familiar with the book. In the IDS condition, the child sat on his

[18] We use the term "Mandarin Chinese" in reference to "Putonghua" or "Standard Chinese", the official language spoken in China. It should be distinguished from Taiwanese Mandarin, another variety of Mandarin Chinese spoken in Taiwan.

or her mother's lap, and the mother was instructed to read the story to her child the way she usually did at home. The mothers were specifically told they could use any sentences; the only requirement was to include the words on each page. In the ADS condition, the mothers were instructed to tell the story to the experimenter (female, a native speaker of Mandarin Chinese), taking into account that she was a college student. This was done to control the speech context and content in both conditions. The order of the two conditions was counterbalanced across participants. A ZOOM H1 recorder (with 16-bit resolution and a sampling rate of 44.1 kHz) was used to make audio recordings, and all sessions were videotaped. Each experimental session took about 15–20 min. All families received a book as a gift after the session.

4.3 Data analysis and results

4.3.1 Data Analysis

The beginnings and endings of the target syllables (the second syllable of each target word) were annotated and extracted from the recordings in PRAAT (Boersma & Weenink, 2017), following the phonetic segmentation principles in Skarnitzl and Machač (2011). In total, 713 target syllables were extracted; of these, 47 syllables (6.6%) were excluded due to background noise or interference from a child's voice.

We chose to acquire the maximum and minimum F0 for each syllable by marking them manually, rather than limiting tone measures to any specific segment(s) within the syllables. This was done for two reasons. First, the domain of tones (or Tone Bearing Units (TBUs)) is phonologically determined, and what constitutes a TBU in Mandarin Chinese is debatable (see Zhang, 2014, p. 81 for a review). Phonetic studies have shown that the voiced parts of syllables—i.e., vowels, initial voiced consonants, prenuclear onglides, and nasal codas—may convey tonal information (Duanmu, 2007; Howie, 1974). In studies involving acoustic analyses of lexical tones in IDS, however, the common practice has been to identify tones based on the F0 measures on vowels, potentially leading to the exclusion of other segments that may carry pitch contours. Second, contextual tonal variations in natural speech—for example, anticipatory and carry-over effects in adjacent Mandarin tones (Xu, 1997)—may also make it difficult to extract pitch measures accurately using an automatic method. In previous studies, the stimuli were either monosyllabic (Xu Rattanasone et al., 2013) or associated with the first syllable of the target words in natural speech, where the carry-over effects from the pre-target syllables were uncertain (Liu et al., 2007). Such methods disregard the potential for contextual impact from adjacent tones. In the current study, we made sure that the first syllable of the target words was always Tone 2 (a rising tone), so that the first syllable had a similar effect on the second tone for each target word.

Taking these issues into account, to get a more accurate picture of tonal information, the author manually marked the maximum F0 and minimum F0 following the methods from Chen and Gussenhoven (2008). As a secondary cue to tones, durations of syllables were extracted automatically using a Praat script (Lennes, 2017). Using these techniques, we obtained four dependent measures for each target syllable: Minimum F0, Maximum F0, F0 range (Maximum F0 – Minimum F0), and Duration of syllables (in seconds). Tone 1 was excluded in the F0 range analyses since it is a flat tone, for which the pitch height (not the pitch range) is the major cue.

For all the F0 measures, we followed Liu et al. (2007) and used two scales: (1) Hz, a linear pitch scale that has been used traditionally in phonetic research; and (2) Equivalent-rectangular-bandwidth-rate (ERB), which has been found to better describe pitch perception (Hermes & Van Gestel, 1991).

To understand whether tones differed between (1) ADS and IDS and (2) IDS directed at 18-month-olds and IDS directed at 24-month-olds, we used linear mixed-effects models for all analyses. In the models, we included fixed factors of Age (18-month-old/24-month-old), Condition (ADS/IDS) and Tone (Tones 1, 2, 3, and 4) on these dependent measures: Minimum F0 (in Hz and ERB), Maximum F0 (in Hz and ERB), F0 range (in Hz and ERB, for Tone 2, Tone 3, and Tone 4, excluding Tone 1), and Syllable duration (in seconds), with Participant Number as a random factor, and allowing for random slopes for Condition and Tone (Barr et al., 2013). All dependent measures were square-root transformed from raw data to get a more normalized distribution (indicated by W in Shapiro–Wilk test).

We used the lme4 package (Bates et al., 2015) in the R environment (R Development Core Team, 2018) for all data analyses. For each dependent measure, we took the backward elimination approach, starting with a model that included all fixed effects plus the random factor, and all interactions between them (the most complex model) (Bates et al., 2015).[19] Then, we used the "step" function in the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) to reduce the models by eliminating non-significant factors or interactions. When we arrived at an interaction of the fixed effects Condition and Age in the final models, we split the data by Age and built further models for each age group.[20] For Maximum F0 (Hz) and Minimum F0 (ERB), the models with maximal random effects failed to converge.

[19] An example of the R codes for these models is: sqrt(max_hz) ~ Condition * Tone * Age + (1 + Condition + Tone | Participant)

[20] An example of the R codes is: sqrt(max_hz) ~ Condition * Tone + (1 + Condition + Tone | Participant)

Therefore, we excluded Tone as a random effect[21] for these two measures. As the results were consistent across Hz and ERB for all pitch measures, we only present results in Hz here. The results of F0 measures in ERB can be found in **Appendix B** (Supplementary materials for Chapter 4). In the following subsections, we report on the final models for each dependent measure. Our main aim was to investigate the general tone hyperarticulation phenomenon in Mandarin Chinese IDS; hence we focus on the fixed effects of Condition and Age, as well as the interaction between these two factors. To further explore whether tonal contrasts are enhanced in IDS, we present an exploratory analysis on the enhancement of Tone 1–Tone 4 contrast in Section "Exploring the Enhancement of Tone 1–Tone 4 Contrast."

4.3.2 Results

4.3.2.1 General tone hyperarticulation

The means and standard deviations for each tone measurements are shown in **Table 4.2**. For Maximum F0 (Hz) (**Figure 4.1**), the final model (**Table 4.3**) revealed a significant main effect of Condition (p = 0.001), as well as a significant interaction of Condition and Age (p = 0.015). To further examine the different effects of Condition on Maximum F0 in the two age groups, we split the data by Age. The models for the two age groups showed a significant main effect of Condition for the 18-month group (β = 1.401, SE = 0.398, t = 3.516, p = 0.002), but not for the 24-month group, suggesting that there was no effect of Condition on Maximum F0 for IDS directed at 24-month-olds. The final models for Maximum F0 and Minimum F0 for each age group can be found in **Appendix B** (Supplementary materials for Chapter 4). Thus, the Maximum F0 of lexical tones was higher in IDS than in ADS only in the 18-month-old group. By the time children were 24 months old, there was no difference between the two speech registers with respect to Maximum F0.

[21] An example of the R codes is: sqrt(max_hz) ~ Tone + (1 + Condition | Participant)

Table 4.2. Mean maximum F0 (Hz), minimum F0 (Hz), F0 range (Hz), and syllable duration (s) of tones, standard deviations in parentheses.

Age	Condition	Tone	Maximum F0 (Hz)	Minimum F0 (Hz)	F0 range (Hz)	Syllable duration (s)
18m	ADS	Tone 1	295.08 (63.4)	264.09 (57.25)	29.70 (22.24)	0.272 (0.131)
		Tone 2	282.10 (81.3)	209.78 (67.55)	72.32 (71.56)	0.302 (0.141)
		Tone 3	269.17 (81.5)	191.70 (49.61)	77.47 (60.47)	0.223 (0.146)
		Tone 4	308.01 (68.6)	228.86 (53.71)	79.16 (66.42)	0.293 (0.139)
	IDS	Tone 1	325.90 (96.2)	287.05 (81.23)	38.85 (36.96)	0.254 (0.121)
		Tone 2	334.81 (125.9)	222.15 (64.87)	112.66 (130.96)	0.322 (0.160)
		Tone 3	300.37 (90.2)	192.45 (51.03)	107.92 (92.06)	0.271 (0.169)
		Tone 4	384.98 (99.1)	262.35 (79.94)	122.63 (111.38)	0.306 (0.143)
24m	ADS	Tone 1	319.31 (91.2)	279.14 (60.14)	40.17 (49.74)	0.242 (0.080)
		Tone 2	283.72 (87.0)	220.61 (61.11)	63.11 (47.87)	0.307 (0.134)
		Tone 3	299.24 (88.3)	198.40 (64.73)	100.83 (71.43)	0.223 (0.119)
		Tone 4	348.89 (94.0)	238.00 (73.57)	110.89 (79.00)	0.259 (0.090)
	IDS	Tone 1	312.23 (106.0)	269.78 (80.95)	42.46 (46.16)	0.256 (0.129)
		Tone 2	274.58 (80.6)	209.82 (53.88)	64.76 (56.03)	0.353 (0.179)
		Tone 3	286.21 (101.0)	184.49 (63.01)	101.72 (83.78)	0.272 (0.185)
		Tone 4	364.17 (126.0)	228.63 (76.11)	135.55 (98.42)	0.283 (0.137)

Figure 4.1. Box plots of Maximum F0 (Hz) for ADS and IDS addressing 18-month-old and 24-month-old children.[22]

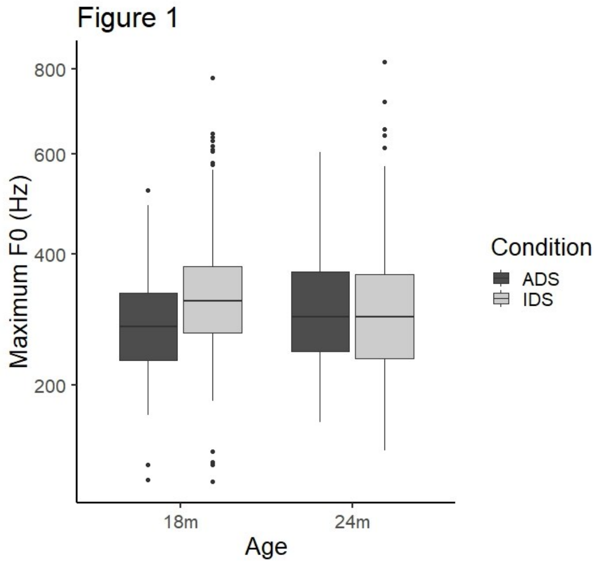


Table 4.3. Final model for Maximum F0 (Hz).

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	16.8043	0.401	41.942	<0.001***
Condition (IDS)	1.403	0.400	3.518	0.001**
Tone2	-0.483	0.259	-1.866	0.063
Tone3	-0.893	0.253	-3.532	<0.001***
Tone4	1.130	0.246	4.600	<0.001***
Age (24m)	0.647	0.538	1.202	0.237
Condition(IDS):Age(24m)	-1.480	0.582	-2.545	0.015*
	Random factors			
	Variance	***SD***		
Participant(Intercept)	2.044	1.430		
Condition (IDS)	1.941	1.393		
Residual	5.040	2.245		

*Note: Intercept represents ADS, Tone 1, and 18 months. * p < 0.05; ** p < 0.01; *** p < 0.001*

[22] All Y-axes in the box plots in this chapter are square-root transformed as the dependent measures were square-root transformed.

Results for Minimum F0 (Hz) (**Figure 4.2**) showed a similar pattern: the final model (**Table 4.4**) showed a significant main effect of Condition (p = 0.030) and a significant interaction of Condition and Age (p = 0.014). When we split the data by Age, we found that there was a significant main effect of Condition for the 18-month-old group (β = 0.589, SE = 0.224, t = 2.630, p = 0.010), but not for the 24-month-old group, as Condition was not in the final model. The results reveal that, similar to Maximum F0, Minimum F0 was also significantly higher in IDS addressing 18-month-old children than in ADS, while no similar differences in Minimum F0 arose between ADS and IDS for the 24-month-old group.

Figure 4.2. Box plots of Minimum F0 (Hz) for ADS and IDS addressing 18-month-old and 24-month-old children.

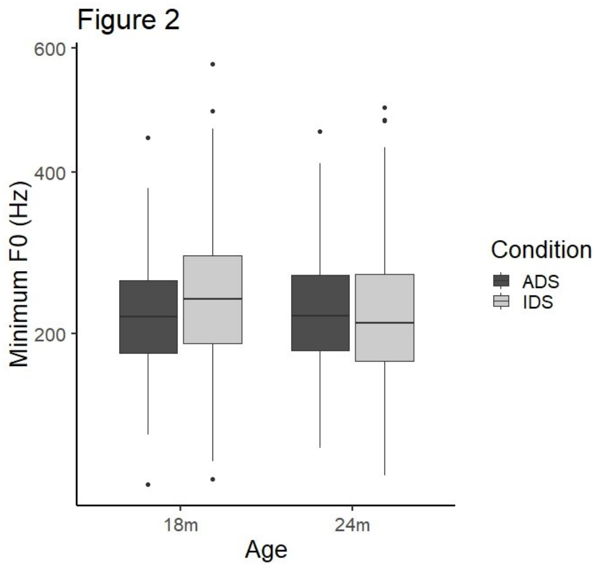


Table 4.4. Final model for Minimum F0 (Hz).

Parameters	***Estimate***	***SE***	***t-value***	***p***
Fixed factors				
(Intercept)	16.214	0.334	48.600	<0.001***
Condition (IDS)	0.552	0.248	2.225	0.030*
Tone2	-1.981	0.300	-6.608	<0.001***
Tone3	-2.970	0.274	-10.825	<0.001***
Tone4	-1.161	0.252	-4.610	<0.001***
Age (24m)	0.390	0.408	0.954	0.346
Condition(IDS):Age(24m)	-0.897	0.352	-2.548	0.014*
	Random Factors			
	Variance	***SD***		
Participant(Intercept)	1.656	1.287		
Condition (IDS)	0.408	0.639		
Tone (Tone 2)	1.592	1.261		
Tone (Tone 3)	1.093	1.046		
Tone (Tone 4)	0.771	0.878		
Residual	3.472	1.864		

Note: Intercept represents ADS, Tone 1, and 18 months. $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

For the measure of F0 range (**Figure 4.3**), the final model (**Table 4.5**) only showed a significant main effect of Condition ($p = 0.006$); no interaction between Age and Condition was observed on this measure, suggesting that lexical tones in Mandarin Chinese IDS have a larger F0 range than ADS tones across the two age groups.

The last measure was duration (**Figure 4.4**). For this measure, the final model (**Table 4.6**) did not include Condition, suggesting that there was no effect of Condition on duration for either the 18-month-old or the 24-month-old groups.

Table 4.5. Final model for F0 range (Hz).

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	8.201	0.576	14.227	<0.001***
Condition (IDS)	0.969	0.352	2.751	0.006*
Tone3	0.050	0.632	0.079	0.937
Tone4	0.480	0.600	0.800	0.424
Age (24m)	-1.288	0.764	-1.686	0.094
Tone3:Age (24m)	1.848	0.884	2.091	0.037*
Tone4:Age (24m)	2.467	0.855	2.884	0.004**
	Random factors			
	Variance	***SD***		
Participant(Intercept)	1.687	1.299		
Residual	14.899	3.860		

*Note: Intercept represents ADS, Tone 1, and 18 months. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$*

Figure 4.3. Box plots of F0 range (Hz) for ADS and IDS addressing 18-month-old and 24-month-old children.

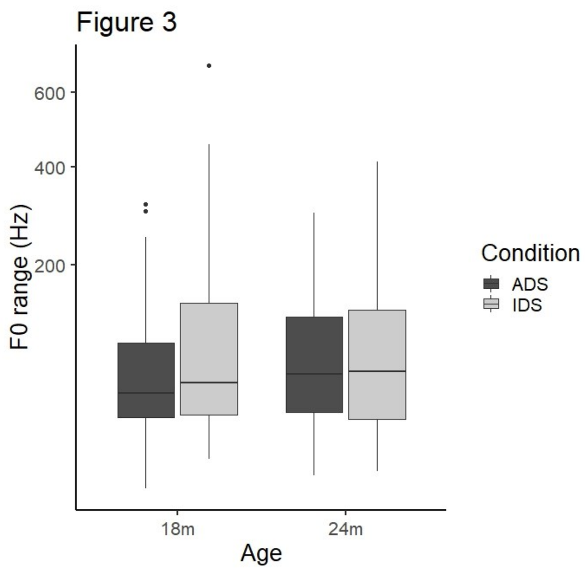
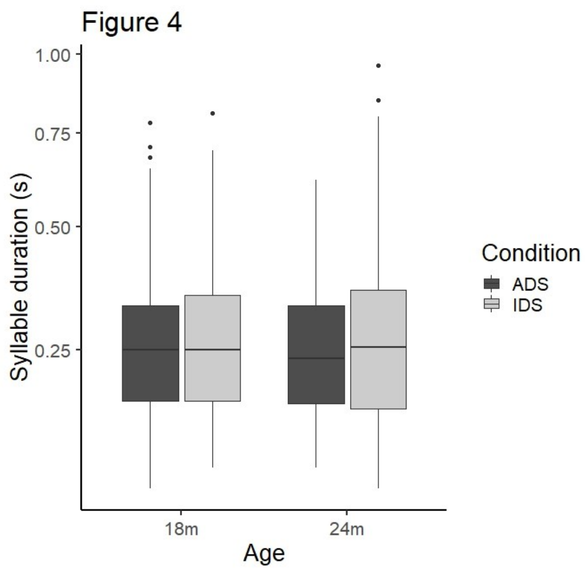


Table 4.6. Final model for Syllable duration (s).

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	0.493	0.011	44.314	<0.001***
Tone2	0.070	0.014	5.230	<0.001***
Tone3	-0.003	0.015	-0.197	0.844
Tone4	0.028	0.013	2.137	0.034*
	Random Factors			
	Variance	***SD***		
Participant(Intercept)	0.004	0.060		
Condition (IDS)	0.004	0.062		
Tone (Tone 2)	0.001	0.025		
Tone (Tone 3)	0.002	0.043		
Tone (Tone 4)	0.001	0.032		
Residual	0.013	0.113		

Note: Intercept represents Tone 1. $p < 0.05$*; ** $p < 0.01$; *** $p < 0.001$*

Figure 4.4. Box plots of syllable duration (s) for ADS and IDS addressing 18-month-old and 24-month-old children.



4.3.2.2 Exploring the enhancement of the Tone 1–Tone 4 contrast

Our main goal was to provide a global measure of tone hyperarticulation, however, tone hyperarticulation may also suggest that tonal contrasts are enhanced in IDS. In addition to comparing the tonal cues between ADS and IDS, we explored whether the contrast between Tone 1 and Tone 4 was enhanced in IDS.[23] Both the contrast between Tone 1 (high-level tone) and Tone 4 (high-falling tone) and the contrast between Tone 2 (mid-rising tone) and Tone 3 (low-dipping tone) are typically used in studies on infant tone perception (e.g., Chen et al., 2017; Liu & Kager, 2017). As the realization of Tone 3 has a large degree of variation in spontaneous speech depending on various factors (e.g., Tone 3 sandhi and the position of a Tone 3 syllable in an utterance, see Yip, 2002), it is impossible to gauge the enhancement of Tones 2–3 contrast from the current data. Thus, we opted instead to focus on the tonal contrast between Tones 1 and 4 and explored whether this tonal contrast was enhanced in IDS as compared with ADS. Since Tones 1 and 4 are mainly distinguished by pitch range, if the difference in pitch range between Tones 1 and 4 was larger in IDS than in ADS, we can conclude that the contrast between Tones 1 and 4 was enhanced in IDS.

First, we took all occurrences of Tones 1 and 4 across the two age groups into analysis. A paired samples t-test showed that there was a marginally significant difference ($t = 2.024$, df = 35, $p = 0.051$) in the difference in pitch range (Tones 4 – Tone 1) between IDS (mean = 97.724 Hz, $SD = 75.934$) and ADS (mean = 66.987 Hz, $SD = 60.927$). As in Liu et al. (2007), we then further considered the first two occurrences of each tone. A paired samples t-test showed that there was a significant difference ($t = 2.294$, df = 31, $p = 0.029$) in the difference in pitch range (Tones 4 – 1) between the two conditions (ADS: mean = 65.408 Hz, $SD = 55.525$; IDS: mean = 103.397 Hz, $SD = 88.669$) as compared with ADS. Taken together these results showed that the contrast between Tones 1 and 4 was enhanced in IDS, especially for the first two occurrences of the target syllables.

4.3.3 Results summary

Both Minimum F0 and Maximum F0 of lexical tones were higher (in both Hz and ERB) in IDS addressing 18-month-old children than in ADS, but no similar differences were observed between ADS and IDS addressing 24-month-children. This pattern suggests that mothers in the study raised the pitch level of tones when they addressed 18-month-old children, but maintained ADS-like pitch height when

[23] It should be noted that the results reported in 4.3.2.2 are exploratory. Our tokens per tone were few resulting in low statistical power for this specific analysis. An alternative way of measuring tonal contrast enhancement would be to compare the tone space area of selected tones (Tone 1, Tone 2, and Tone 4) in ADS and IDS as in Tang et al. (2017). Tang and colleagues found that tones in Mandarin Chinese IDS were only hyperarticulated at the utterance-final positions. However, in our current data set, only 18 out of 39 participants had tones (Tone 1, Tone 2, and Tone 4) produced at utterance-final positions in both ADS and IDS conditions, meaning that tone space area could only be compared for these participants. As a result, such an analysis is not feasible for our data set.

addressing 24-month-olds. F0 range (Hz and ERB), on the other hand, showed a difference between ADS and IDS across ages: F0 range was larger in IDS compared with ADS for both 18- and 24-month-olds. As for duration, our results showed that tones were not lengthened in either age group.

Our results showed that tone hyperarticulation was present in IDS addressing 18- and 24-month-old children, but the specific tonal cues differed between the two groups: for 18-month-olds, Tone 1 had a higher F0 in IDS, and Tones 2, 3, and 4 had higher F0 and a larger F0 range in IDS. For 24-month-olds, all four tones remained the same pitch level in the two speech registers, though Tones 2, 3, and 4 in IDS still had a larger pitch range in IDS. As a secondary cue to lexical tone (Blicher et al., 1990), duration did not differ between ADS and IDS in either age group. In addition, an exploratory analysis showed that the contrast between Tones 1 and 4 was enhanced in IDS.

4.4 Discussion and conclusions

This study examined lexical tones in Mandarin Chinese IDS addressing 18- and 24-month-old children, at the age of the vocabulary spurt. The study had two main goals: to test whether tones are hyperarticulated in IDS compared with ADS in Mandarin Chinese, and to explore how tones in IDS vary with the age of the addressee during the period of vocabulary spurt. To accomplish these goals, we measured the acoustic cues of lexical tones in ADS and IDS in a semi-spontaneous story-telling task. The results demonstrated that tone hyperarticulation and age-related changes are observed in Mandarin Chinese IDS addressing toddlers.

Our research questions were: (1) Are tones in Mandarin IDS addressed to 18- and 24-month-old children hyperarticulated compared to tones in ADS? (2) Does lexical tonal hyperarticulation in Mandarin Chinese IDS change when the mother is addressing an 18-month-old child vs. a 24-month-old child? Our results build on past studies on lexical tones in IDS addressing preverbal children (Liu et al., 2007; Xu Rattanasone et al., 2013), while extending that research to the second year of life—the period of the vocabulary spurt. Our findings show that tone hyperarticulation remains present in speech to toddlers, even after their phonetic perception has tuned to their native language and they have started learning words. Specifically, we found that, in speech addressed to 18-month-old children, both the minimum and maximum F0 of tones was higher in IDS than ADS, and the F0 range was larger, but the tones were not lengthened. These F0 measures are consistent with the findings of Liu et al. (2007) for IDS addressing 12-month-old children. In speech addressed to 24-month-old children, we found that pitch height of lexical tones had normalized to the ADS standard, while F0 range remained larger in IDS than ADS. Tone duration does not appear to differ between toddler-addressed IDS and ADS.

Taken in the context of previous studies exploring lexical tones in IDS addressing preverbal children and preschool children, our results contribute to the timeline of tonal changes in IDS by providing evidence for tone hyperarticulation in the second year. Xu Rattanasone et al. (2013) demonstrated that tones in Cantonese IDS are hyperarticulated when talking to children from 3 to 9 months, but that hyperarticulation declines as children approach 12 months. In their study on Taiwanese Mandarin, Liu et al. (2007) found that when addressing 12-month-old children, mothers exaggerated every acoustic correlate of tone in IDS, including producing a higher F0, larger F0 range, and longer duration. The current findings fill a crucial gap in the timeline and suggest that tone hyperarticulation may continue until children reach their second birthday. Liu et al. (2009) compared tones in Taiwanese Mandarin–speaking mothers' speech to preverbal children (IDS, age range: 0;7–1;0) and speech to preschool children (CDS: age: 5;0), and found that the degree of tone exaggeration was much less in CDS than IDS. Based on this evidence, we may tentatively trace a developmental trajectory of tone hyperarticulation in IDS: hyperarticulation is notably salient from birth to 12 months in both F0 and duration measures, remains present for F0 measures of tone at 18 months, but begins normalizing toward the ADS standard by the end of the second year. By 24 months, the degree of pitch height difference between ADS and IDS drops significantly for all four tones, although pitch range (of Tones 2, 3, and 4) remains larger in IDS compared with ADS. However, simply combining these findings is not sufficient to produce a complete picture of the developmental trajectory of how lexical tones change with age in IDS, since the studies noted above investigated different tone languages, and adopted different acoustic measures and different elicitation methods.

A question that follows from these findings is: why do tonal cues in IDS change over time? It seems likely that the change in the pitch level (minimum and maximum F0) is related to the general prosodic exaggeration, as the degree of prosodic exaggeration in IDS may also decline from 18 to 24 months when children have become more verbal and their word learning accelerates. However, since studies on tone hyperarticulation (including the current study) usually focus on the syllabic level, little is known about whether tonal cues coincide with other prosodic features of IDS. Crucially, our results showed that the pitch range (of Tones 2, 3, and 4) remained enlarged in IDS even when the pitch height had declined to the ADS level at 24 months, suggesting that mothers may hyperarticulate lexical tones during the period of vocabulary spurt in support of word learning.

A relationship between the quality of IDS and children's language development has often been assumed in research on IDS (e.g., Fernald & Simon, 1984), and the hyperarticulation phenomenon has been offered up as evidence for the facilitative effects of IDS on language acquisition. However, although phonetic input is clearly exaggerated in IDS, at the same time, it is also highly variable compared with the input observed in ADS (Adriaans & Swingley, 2017). Might this variability make it more difficult for infants to form phonetic categories? Adriaans and Swingley's (2017) research suggests not: the authors used categorization models to train two datasets of hyperarticulated vowels (IDS-characterized) and non-exaggerated vowels (ADS-characterized), and found that the highly variable vowels in IDS favored phonetic categorization compared with the non-exaggerated vowels. However, empirical research on whether IDS indeed supports language acquisition—and more specifically, tone acquisition and word learning in tone languages—is surprisingly lacking. Future research should examine whether raised pitch and/or enlarged pitch range indeed facilitates children's word recognition and word learning.

Thus, we must be cautious in interpreting our results as direct evidence for the linguistic function of IDS in word learning. Indeed, although the current study demonstrates that tone hyperarticulation remains present in language input during the vocabulary spurt period, it does not necessarily indicate that children benefit from this linguistic phenomenon. Several studies have explored the correlation between the quality of IDS and children's language outcomes. For instance, Liu et al. (2003) found that the vowel space in Taiwanese Mandarin IDS toward preverbal children (6–8 months; 10–12 months) is related to infants' performance on speech discrimination. Hartman et al. (2017) further showed that the quality of vowels in early English IDS may predict vocabulary size among 2-year-olds. In a word-learning study, Ma et al. (2011) found that 21-month-old English-learning children could only learn words in the IDS condition, while 27-month-old children could learn words successfully in both IDS and ADS conditions. For tone languages, the correlated question—whether tone hyperarticulation in IDS indeed benefits lexical word learning—remains under-investigated. At this point, no research exists directly comparing word learning under ADS and IDS conditions in tone languages, and the literature offers no insight into how tones in language input correlate with vocabulary outcomes.

The pitch measures of tones addressed to our 24-month-old group showed a different pattern from the findings in Liu et al. (2007). In addition to the different age groups under investigation, an alternative explanation for the inconsistent results may be attributed to language-specific properties. Even though Mandarin Chinese (spoken in mainland China) and Taiwanese Mandarin are variations of the same language, their sentential prosody differs (Chen et al., 2009), which may in turn affect the prosody of IDS. Literature comparing the prosody of IDS in Taiwanese Mandarin and Mandarin Chinese is lacking. As British English and American English IDS exhibit different prosodic features (Fernald et al., 1989), one direction invited by the current research is to compare tone hyperarticulation in different tone languages, as well as different variations of the same tone language.

A limitation of the current design is that we used only one target word for each tone. We also took steps to avoid generating contrasts between the target words by ensuring that the phonemes of the target syllables differed from each other. As vowels and tones may interact (Hoole & Hu, 2004), our results are not generalizable to all syllable–tone combinations in Mandarin Chinese. However, it should be noted that this line of research typically relies on a rather small set of stimuli due to the practicalities of testing children. For example, Kuhl et al. (1997) used one target word per vowel; Liu et al. (2007) had twelve syllables for four tones, but they only included "the first two clear tokens of each target word"; Tang et al. (2017) also used one syllable for each tone. Even though there has been some agreement on the salience of tone hyperarticulation in the first year of life, it has not been established that tone hyperarticulation is present across the board. Meta-analysis of existing tone hyperarticulation studies may provide a better understanding of this issue. Also, the current study took a cross-sectional design. We found no effect of Age across ADS and IDS, indicating that there are no group differences between the 18-and 24-month-old groups. However, a timetable of changes in tone hyperarticulation over time remains to be revealed by longitudinal studies.

Another useful future direction for study would be to examine whether tone hyperarticulation is related to the prosodic marking of focused words. Previous research has shown that, in English IDS, mothers tend to put contextually new words (focused words) at utterance-final positions, and these focused words usually carry prosodic marking in the form of higher pitch and a larger pitch range (Fernald & Mazzie, 1991). Relatedly, Tang et al. (2017) showed that tone hyperarticulation in Northern Mandarin only occurs in utterance-final position. In their experimental design, toys corresponding to the target words were provided one by one to each participant (thus, each target word was contextually new). As Mandarin Chinese and English are both SVO languages, it is certainly possible that tone hyperarticulation in Northern Mandarin, as in English, tends to occur when the lexical item in question is the focus of an utterance.

This study investigated the tone hyperarticulation phenomenon in Mandarin Chinese IDS in the second year of life and revealed age-related changes of tonal cues in IDS addressed to 18-month-old vs. 24-month-old children. These findings may contribute to an understanding of the role of IDS in tone acquisition and word learning. Mothers may hyperarticulate lexical tones in order to provide more fine-grained information for language acquisition. However, it may be premature to interpret these findings as direct evidence for the linguistic function of IDS.

Chapter 5

Is prosody of infant-directed speech in word-learning contexts correlated with children's vocabulary? Evidence from Dutch and Mandarin Chinese

Abstract

The most salient feature of infant-directed speech (IDS) is its exaggerated prosody compared to adult-directed speech (ADS). While the prosodic features of prototypical IDS have been shown to attract infants' attention and convey positive affect, whether IDS prosody has specifically linguistic functions is still a matter of debate. In particular, it remains unclear whether there are correlations between the prosody of IDS and children's vocabulary size. Also, no study has considered the prosody of IDS in word-learning contexts in which mothers introduce unfamiliar words to children as a predictor. This chapter set out to examine the issue by asking whether the global prosody of IDS and/or the prosody of IDS specific to word-learning contexts are correlated with children's vocabulary size. We collected speech data of Dutch and Mandarin Chinese IDS in a semi-spontaneous storybook-reading task in which mothers introduced familiar or unfamiliar words to their 18- and 24-month-old children. Multiple regressions revealed that for Dutch, the prosody of IDS in word-learning contexts, rather than the generally exaggerated prosody, was correlated with children's vocabulary size as well as children's vocabulary growth from 18 to 24 months. However, no significant correlations were found for Mandarin Chinese. Together these findings suggest that the prosody of (Dutch) IDS specific to word-learning contexts may serve linguistic purposes.

5.1 Introduction

Children learn words from linguistic input, and it is well established that the quantity and quality of language input a child receives are associated with his or her lexical development (e.g., Hart & Risley, 1995; Hoff & Naigles, 2002; Ramírez-Esparza, García-Sierra, & Kuhl, 2014). Infant-directed speech (IDS) is a speech register mothers typically use when addressing their child. The major characteristic that distinguishes prototypical IDS from adult-directed speech (ADS) is its exaggerated prosody (Cristia, 2013). While there is robust evidence to suggest that the exaggerated prosody of IDS attracts infants' attention and regulates infants' emotion (see a review in Spinelli et al., 2017), whether the exaggerated prosody has specifically linguistic functions is still debatable. So far, a few studies have shown that some prosodic features (e.g., pitch range and speaking rate) of IDS are associated with children's vocabulary size (e.g., Porritt et al., 2014; Raneri, 2015), while others suggest the opposite (e.g., Kalashnikova & Burnham, 2018). These correlational studies have invariably measured prosody on a global level, rather than focusing on the prosodic input specific in word-learning contexts. When talking to their child, mothers often need to introduce words that are novel or unfamiliar to the child. In a typical word-learning context during mother-child interaction, mothers refer to a novel object with a novel word label, while their child learns to associate the novel label to the novel object. Even if the generally exaggerated prosody of IDS may not be reliably associated with children's vocabulary size, the prosody of IDS in word-learning contexts may be related to children's vocabulary size. The goal of this study is to investigate the links between IDS prosody and lexical development. In particular, we examined (1) whether the generally exaggerated prosody (temporal and pitch properties) of IDS could predict children's vocabulary size; and (2) whether the IDS prosody (temporal and pitch properties) specific to word-learning contexts is associated with children's vocabulary size. To achieve this goal, we examined two languages that differ in their prosodic characteristics: Dutch, a stress-timed and stressed language and Mandarin Chinese, a syllable-timed and tonal language.

5.1.1 Infant-directed speech and lexical development

In the literature, the term "infant-directed speech" (IDS) may have one of two interpretations. First, "IDS" typically refers to a speech register that caregivers use when addressing children. Compared to ADS, prototypical IDS is characterized by a higher pitch, a larger pitch range, and a slower speaking rate compared to ADS (see reviews in Cristia, 2013; Soderstrom, 2007). This speech register is used by caregivers in almost all languages and cultures across the world (Fernald et al., 1989; Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002; Narayan & McDermott, 2016).

The addressers of IDS include but are not limited to mothers, fathers, and grandmothers (Fernald et al., 1989; Shute & Wheldall, 2001). As in most studies, however, we focus on mothers' IDS in this study.

Second, in some of the literature, the term IDS can also refer to *any* speech that is directed to a child (e.g., Kalashinikova & Burnham, 2018). Contrary to the common assumption that all speech addressed to children is modified in prosody, speech directed to children shows great variation and may not always be exaggerated in prosody to the same extent (to be discussed later). Thus, in this study we use the term "prototypical IDS" when we highlight its prosodic differences from ADS in order to avoid ambiguity.

Correspondingly, two approaches have been taken here to examine the role of prosodic input in children's early language development. First, as prototypical IDS has long been claimed to facilitate language acquisition as compared to ADS (e.g., Fernald, 1985; Kuhl et al., 1997), studies have compared children's word learning performance online when children hear either ADS or prototypical IDS. Second, studies have used mothers' individual IDS prosody as predictors for their child's vocabulary size in order to assess whether the prosodic quality of IDS is correlated with children's vocabulary size. We review the two approaches in the following paragraphs.

First, prototypical IDS facilitates children's online performances related to word learning, including word segmentation, word recognition, and word-to-object mapping. In these studies, children's online word learning performance is compared between ADS and IDS conditions. Children hear auditory stimuli that have similar speech content but are produced with either ADS or prototypical IDS prosody: IDS stimuli usually have a higher pitch, larger pitch range, and a slower speaking rate compared to ADS stimuli. If children's online learning performances are better in the IDS condition compared to the ADS condition, the results are interpreted as evidence for the facilitative effects of IDS. For example, English- and German- learning infants (7 months old) could only segment words in continuous speech when hearing prototypical IDS but not when hearing ADS input (Thiessen, Hill, & Saffran, 2005; Mani & Pätzold, 2016). English-learning infants were able to recognize words that they had been familiarized with in prototypical IDS even after 24 hours, but not when the words were introduced in ADS (Singh et al., 2009). Regarding word-to-object mapping, Graf Estes and Hurley (2013) found that 17.5-month-old English-learning children only learned word-object pairings when the words were produced in prototypical IDS, but they failed to learn words in the ADS condition. Similar results are found for older children: English-learning 21-month-old children succeeded at word-to-object mapping when listening to the auditory forms of words presented in prototypical IDS but not in ADS. Only after 27 months could children reliably learn novel words presented in ADS (Ma et al., 2011). Together these results suggest that the prototypical IDS prosody facilitates online word learning compared to ADS.

Even though the studies illustrated above show robust evidence for the facilitative effects of prototypical IDS on online language processing, it should be noted that IDS in natural mother-child interactions varies to a great extent among individuals. Such individual variations exist in both quantity and quality of IDS. For example, Hart and Risley (1995) found that the quantity of input differs to a considerable degree between high socioeconomic status (SES) and lower SES families. Regarding prosodic quality, Narayan and McDermott (2016) examined IDS in Korean, Tagalog, and Sri Lankan Tamil and found that while mothers modified pitch and speaking rate in IDS compared to ADS in general, some mothers showed no evidence of pitch or temporal modifications in their speech to their child. Due to individual variations, even though prototypical IDS may enhance children's language learning under experimental conditions in labs, the first approach fails to demonstrate the link between individual IDS and children's vocabulary size.

Several studies therefore took the second approach which explores the correlations between prosodic properties in IDS and children's vocabulary size. Despite extensive literature on the correlations between input quantity and input quality and children's vocabulary size, only a few studies have investigated whether the quality of *prosodic* input is correlated with children's vocabulary size (e.g., Kalashnikova & Burnham, 2018, to review later).

With respect to input quantity, robust evidence shows that children who hear more words in parent-child interactions tend to develop a larger vocabulary (Hart & Risley, 1995; Hoff, 2006; Hoff & Naigles, 2002; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Mahr & Edwards, 2018). Also, there is some evidence to suggest that the quantity of prototypical (English) IDS (which has an exaggerated prosody) in a child's input is specifically associated with his or her vocabulary size (Ramírez-Esparza et al., 2014).

In addition to input quantity, many (non-prosodic) aspects of input quality, such as lexical richness, syntactic complexity, repetitiveness, and vowel hyperarticulation are all found to be correlated with individual child's vocabulary size (Hartman, Bernstein Ratner, & Newman, 2016; Hoff & Naigles, 2002; Kalashnikova & Burnham, 2018; Liu, Kuhl, and Tsao, 2003; Newman, Rowe, & Bernstein Ratner, 2015).

Even though prosodic features—pitch modifications and temporal modifications in IDS—are the most prominent characteristics of IDS, only a few studies have investigated the relationship between IDS prosody and children's vocabulary size. For instance, Porritt, Zinser, Bachorowski, and Kaplan (2014) examined the relationship between pitch properties in English IDS and children's vocabulary size. Their pitch measures of IDS include mean F0, F0 range, and F0 variability (measured by the dispersion of F0 in relation to mean F0). They found that infants' expressive vocabulary percentile scores were positively correlated with F0 range in IDS. In addition to pitch cues, one study showed that the temporal properties of American English IDS—speaking rate—is related to vocabulary size (Raneri, 2015). Specifically, slow speaking rate in IDS at seven months predicts larger expressive vocabulary at two years of age.

Recently, Kalashnikova and Burnham (2018) investigated whether three factors—vowel hyperarticulation, pitch, and affect—in IDS predicted children's vocabulary size at later ages. They measured vowel triangle areas, mean F0, and affect scores (rated by native speakers) in IDS addressed to children at 7, 9, 11, 15, and 19 months of age, as well as in ADS. For each of the three factors, a hyper-score was obtained by dividing each mother's IDS score by their corresponding ADS score. These hyper-scores indicate how exaggerated individual IDS was for each participant mother. That is, if the hyper-score is larger than 1, then IDS was exaggerated compared to ADS for that participant. Furthermore, larger hyper-scores indicate a more exaggerated IDS prosody. The results showed that in all three factors at all ages, IDS was exaggerated compared to ADS. The hyper-scores were then used as predictors for children's vocabulary size. They found that only vowel hyperarticulation hyper-scores at 9 months and beyond were significantly correlated with children's expressive vocabulary size at 15 and 19 months. However, neither pitch nor affect hyper-scores could predict children's vocabulary size. The authors conclude that vowel hyperarticulation, but not generally exaggerated pitch or positive affect, plays a role in lexical development. Similarly, Song, Demuth, and Morgan (2018) did not find any significant correlations between the prosody of individual mothers' IDS (mean pitch and pitch range) at 17 months and children's vocabulary size at 19 or 25 months.

Taken together, online word learning studies suggest that the prototypical prosody of IDS facilitates children's online word learning. Also, at the individual level, the quantity and some aspects of quality in individual IDS input are found to be associated with a child's vocabulary size. However, to date, only a few studies have explored the relationship between IDS prosody and children's vocabulary, and results are inconsistent: Porrit et al. (2014) showed that pitch range was a significant predictor for children's vocabulary, while Kalashnikova and Burnham (2018) and Song et al.

(2018) failed to show a significant effect of mean pitch. Also, only one study to date has provided evidence that the slow speaking rate in IDS may be correlated with children's vocabulary development (Raneri, 2015).

There are two possible accounts for the inconsistent results in the literature reviewed above. One possibility may be attributed to the function of pitch in IDS. It is possible that the function of the generally exaggerated pitch in IDS is mainly attentional, but not linguistic. Spinelli, Fasolo, and Mesman (2017) examined the role of IDS prosody in language acquisition in the first two years of life by conducting meta-analyses of 15 studies on the role of prototypical IDS in language acquisition. Two types of outcomes, attentional and communicative, were included in the meta-analyses. Specifically, the attentional outcomes included any measures of attention, such as global attention and joint attention, and the communicative outcomes included pre-linguistics outcomes, for example infants' response and imitation of input, as well as linguistic outcomes such as vocabulary size, language production and comprehension, etc. Their results suggest that prototypical IDS prosody has a much greater effect on attentional and pre-linguistic aspects such as eliciting vocal responses than it does on linguistic outcomes.

Another possible account for the inconsistent results may lie in the choice of predictors. First, most studies used mean pitch values in the IDS of each mother as the prosodic measure. As such, the individual speaking style of the mothers was not accounted for. One exception is Kalashnikova and Burnham (2018) who used a hyper-score measure by taking ADS as a baseline. By adopting a hyper-score measure, they controlled for the effect of individual variations. Second, the predictors of the studies were all global measures of prosody instead of measurements of prosody in specific interactive contexts such as word-learning contexts. However, IDS varies among different daily activities and contexts even for the same mother-child dyad in natural mother-child interactions. Given that the quantity, diversity, and pragmatic functions of IDS differ by activity (Tamis-LeMonda, Custode, Kuchirko, Escobar, & Lo (2018), differences in IDS are likely to exist with respect to prosody as well. Specifically, none of these studies targeted word-learning contexts in which mothers introduce novel words to children. Such contexts may be assumed to provide the most direct input for children to learn novel words and thus may be considered as crucial input for word learning.

In light of the effect of individual speaking style and the possible variations in contexts, we made two methodological choices. First, we adopted the hyper-score measure from Kalashnikova & Burnham (2018) to reduce individual variations in speaking style. Second, we collected speech data in a word-learning context. Even though Kalashnikova and Burnham (2018) found that the global measures of prosody are not correlated to vocabulary size, it is nonetheless still possible that the use of pitch specifically in word-learning contexts is critical for children's word learning and can be correlated with children's vocabulary size. Thus, in addition to the "general hyper-score" used by Kalashnikova and Burnham (2018), we obtained an "unfamiliar hyper-score" for each prosodic measure to specify how IDS was exaggerated when mothers introduced unfamiliar words to their child.

5.1.2 Cross-linguistic differences

Most studies reviewed above are carried out on English IDS, with only one exception on Spanish (Hurtado et al., 2008 on quantity of input and vocabulary) and one on Taiwanese Mandarin (Liu, Kuhl, & Tsao, 2003 on vowel hyperarticulation and vocabulary). As for the relationship between IDS prosody and vocabulary size, all studies reviewed above investigated English only.

IDS is often considered universal, as almost all languages exaggerate prosody when talking to young children (but see Bernstein Ratner & Pye, 1984 on Quiché Mayan for an exception). However, cross-linguistic variations exist in IDS. Such cross-linguistic differences have been shown with respect to the degree of prosodic exaggeration. For example, for pitch properties, Fernald et al. (1989) found in a cross-linguistic investigation that while IDS prosody was more exaggerated compared to ADS in all languages under investigation, the difference in mean pitch between American English ADS and IDS was larger than in British English, French, Italian, German, and Japanese. Regarding speaking rate, IDS is generally found to be slower than ADS in English, Dutch, German, and many other languages (e.g., Fernald & Simon, 1984; Han, de Jong, & Kager, 2018b; see also Chapter 2 in this dissertation). However, in a cross-linguistic comparison, Narayan and McDermott (2016) found that IDS in Korean and Tagalog were slower in IDS compared to ADS, while Sri Lankan Tamil IDS did not show evidence of slowing down. Also, Mandarin Chinese IDS addressed to 18- and 24-month-old children has not shown evidence of slowing down in comparison to ADS (Han, de Jong, & Kager, 2018b; Chapter 2 of this dissertation). It therefore remains unknown whether the relationships—or lack thereof—between prosodic cues and lexical development as reviewed above can be generalized to languages other than English.

In addition to variations in the degree of prosodic exaggeration, different languages may use different prosodic cues to highlight unfamiliar words in word-learning contexts in IDS (see Chapter 2 and Chapter 3 in this dissertation). Due to the potential differences in the degree of IDS prosodic exaggeration and the different prosodic cues in word-learning contexts, it is possible that the specific prosodic cues that are potentially correlated with vocabulary sizes may differ between these two languages.

5.1.3 The current study

As illustrated above, the results on the relationship between IDS prosody and children's vocabulary size are inconsistent. Also, there is a lack of studies examining the link between prosodic input specific to word-learning contexts and children's vocabulary size. Thus, the current study set out to explore the link between IDS prosody and children's vocabulary size by examining both the generally exaggerated IDS prosody as well as IDS prosody specific to contexts in which mothers introduce words that are unfamiliar to their child. The overarching goal is to determine whether the global prosody of IDS and/or the prosody of IDS specific to word-learning contexts is associated with children's vocabulary size. To achieve this goal, we used two prosodic "hyper-scores" on the global IDS prosody and the word-learning specific IDS prosody, respectively: "general hyper-scores" and "unfamiliar hyper-scores." Following Kalashnikova and Burnham (2018), these hyper-scores were calculated by dividing the IDS values by the ADS values. We predict that the prosodic hyper-scores will be correlated with children's vocabulary size. Specifically, as the prosody in word-learning contexts is immediately relevant to children's word learning, we predict that the prosody of IDS when mothers introduce unfamiliar words to their child (indicated by unfamiliar hyper-scores) will be correlated with children's vocabulary size. Also, we predict that the specific prosodic cues that were related to children's vocabulary size might differ between Dutch and Mandarin Chinese, as these two languages are typologically distinct in prosody.

5.2 Methods

In order to address the research questions, we used the speech data reported in Chapter 2 and Chapter 3 of this dissertation to calculate prosodic hyper-scores. The materials and procedures of two production experiments are described in Chapter 2 and Chapter 3. In the two experiments, we used a storybook-telling task to elicit semi-spontaneous speech from mothers speaking Dutch or Mandarin Chinese. This storybook contained words that were familiar or unfamiliar to children. Each mother told the story twice, once to an adult (ADS condition) and once to their child (IDS condition). After each experiment, mothers completed a checklist to determine the familiarity of words for each child, and the actual familiarity of a word was indicated by this checklist. The procedure of the experiments in Dutch and Mandarin Chinese was similar. The major difference between the two experiments was the design. The Dutch experiment followed a longitudinal design, and mother-child dyads participated when the child was at 18 and 24 months old. The Mandarin Chinese experiment had a cross-sectional design. In addition to the speech data, we collected children's vocabulary scores.

5.2.1 Experiment 1: Dutch

5.2.1.1 Participants

Forty-nine Dutch-speaking mother-child dyads participated when the children were 18 months old (mean age of children = 18;15, age range = 18;00–18;29; girls *N* = 26; mean age of mothers = 35 years, age range = 29–44 years). All children were typically developing with no report of language or hearing problems. The Dutch mother-child dyads were recruited from the Utrecht Baby Lab database and were all Dutch native speakers living in the Utrecht area in the Netherlands. All mothers had higher education. Thirty-two of these participants visited the lab again when the children were 24 months old (mean age of children = 24;18, age range = 24;00–26;30).

5.2.1.2 Vocabulary size

In addition to the speech data, we collected children's vocabulary information using the Dutch adaptation of MacArthur-Bates Communicative Development Inventories (CDI, Fenson, Bates, Dale, Marchman, & Reznick, 2007). All Dutch mothers completed the N-CDI: *Woorden en Zinnen* (Dutch version of the MacArthur Communicative Development Inventory: words and sentences) (Zink & Lejaegere, 2002) online when children were 18 months and 24 months of age. This is a checklist of children's receptive and expressive vocabulary and it contains 702 word items. According to the results of this parental report, at 18 months, children had a mean receptive vocabulary of 247 (*SD* = 103; range = 101–473) and an expressive vocabulary of 65.9 (*SD* = 55.6; range = 2–239). At 24 months, children had a mean receptive vocabulary of 529 (*SD* = 90; range = 352–670) and an expressive vocabulary of 378 (*SD* = 113; range = 188–566).

5.2.2 Experiment 2: Mandarin Chinese

5.2.2.1 Participants

Twenty-one Mandarin-Chinese-speaking mothers of 18-month-old children (mean age = 18;15, age range = 17;21–18;27; girls *N* = 9; mean age of mothers = 30 years, age range = 25 – 39 years) and nineteen mothers of 24-month-old children (mean age = 24;13, age range = 23;2–24;30; girls *N* = 10; mean age of mothers = 31 years, age range = 32–36 years) participated in the study. The Mandarin Chinese dyads were recruited from kindergartens in Yichang, China. All children were typically developing with no report of language or hearing problems. All the participant mothers had higher education and they spoke Mandarin Chinese (Putonghua, the official language in China) proficiently.

5.2.2.2 Vocabulary size

Mandarin Chinese mothers completed the Mandarin Chinese version of CDI (M-CDI, Tardif, Fletcher, Liang, & Kaciroti, 2009). The M-CDI only reports expressive vocabulary and contains 801 word items. The results of the parental report showed that the 18-month-old children had a mean expressive vocabulary of 237 (*SD* = 211, range = 9–786). The 24-month-old children had a mean expressive vocabulary of 542 (*SD* = 152, range = 276–798).

5.2.3 Prosodic measurements

We extracted the following prosodic measures: one temporal measure (articulation rate, measured in syllables/s) and two pitch measures (mean F0 (Hz) and F0 range (st)). The prosodic measures were calculated for each target word as well as for utterances containing the target words (target utterances) (see Chapter 2 and Chapter 3 in this dissertation for details on annotation and data treatment). The prosodic measurements were then averaged by Condition (ADS/IDS) and Familiarity (Familiar/Unfamiliar) for each mother.

Following Kalashnikova and Burnham (2018), for each mother, we calculated hyper-scores for each prosodic measurement by dividing IDS values by the corresponding ADS values. By doing so, we used ADS as a baseline and controlled for individual speaking style. For each measurement, we calculated two hyper-scores: a general hyper-score and an unfamiliar hyper-score.

General hyper-scores in the current study correspond to the hyper-scores in Kalashnikova and Burnham (2018), calculated by dividing IDS by ADS regardless of Familiarity. This set of hyper-scores indicates the general prosodic exaggeration, including measures on the word level and measures on the utterance level. These general hyper-scores indicate how IDS prosody is generally exaggerated compared to ADS regardless of the familiarity of target words. The hyper-scores are: word articulation rate general hyper-score, word mean F0 general hyper-score, word F0 range general hyper-score, utterance articulation rate general hyper-score, utterance mean F0 general hyper-score, and utterance F0 range general hyper-score.

Unfamiliar hyper-scores only included unfamiliar words and utterances with unfamiliar words but not familiar words. This set of hyper-scores shows how exaggerated the unfamiliar words were in IDS relative to ADS in word-learning contexts in which mothers introduce unfamiliar words to their child. The hyper-scores are: word articulation rate unfamiliar hyper-score, word mean F0 unfamiliar hyper-score, word F0 range unfamiliar hyper-score, utterance articulation rate unfamiliar hyper-score, utterance mean F0 unfamiliar hyper-score, and utterance F0 range unfamiliar hyper-score.

5.2.4 Data analysis

For each language, we conducted a series of multiple regression analyses to examine whether the prosodic hyper-scores were correlated to vocabulary size concurrently or longitudinally, and if so, which hyper-score(s) significantly predict children's vocabulary. The multiple regressions were done in the R environment (R Core Team, 2018) using the lm() function. The outcome variables were children's receptive vocabulary or expressive vocabulary at 18 months or 24 months. The predictor variables were hyper-scores. Before building each model, we detected outliers by visual inspection of scatter plots and these outliers were capped at the 5th (for outliers below the lower limit) or the 95th percentile (for outliers above the upper limit). For each model, we started with including all the predictors[24] and then used the "stepAIC" function of the MASS package (Venables & Ripley, 2002) to reduce the model by selecting variables with a significance level of 5% (direction was set to "backward"). We also checked multicollinearity among the predictor variables. If two predictor variables were highly correlated (r > 0.8), we excluded one of the two predictors before we built each model.

5.3 Results

5.3.1 Experiment 1: Dutch

As 49 Dutch participants (girls N = 26) were tested at 18 months, and 32 of the participants were tested again at 24 months (girls N = 19), we performed three sets of multiple regression analyses. (1) Concurrent correlations at 18 months. Specifically, we examined whether there were concurrent correlations between the hyper-scores at 18 months and children's vocabulary size at 18 months. For this analysis, we included speech data from all 49 participants. Six participants were excluded due to missing vocabulary information, resulting in a total number of 43 participants in the final analyses. (2) Concurrent correlations at 24 months. Here we examined whether there were concurrent correlations between the hyper-scores at 24 months and children's vocabulary size at 24 months. For this analysis, we included the 32 participants who participated at both ages, of which 5 participants were excluded due to missing

---

[24] An example of the R codes is: lm(receptive_vocabulary ~ word_AR_genhyper + utterance_AR_genhyper + utterance_mean_genhyper + word_mean_genhyper + utterance_range_genhyper + word_range_genhyper)

vocabulary information at 24 months. Also, as no unfamiliar words were reported for two participants at 24 months, these two participants were excluded from the "unfamiliar hyper-score" analyses at 24 months. (3) Longitudinal correlations between the hyper-scores at 18 months and children's vocabulary at 24 months. In particular, we examined whether there were longitudinal correlations between the hyper-scores at 18 months and children's vocabulary size at 24 months. For this analysis, we also only included the 32 participants who participated at both ages, of which 5 were excluded due to missing vocabulary information at 24 months. For this analysis, the effect of individual differences in vocabulary size was accounted for by including children's vocabulary size at 18 months as a predictor in the model. **Table 5.1** shows the mean general hyper-scores and mean unfamiliar hyper-scores in Dutch.

Table 5.1. Mean general hyper-scores and mean unfamiliar hyper-scores in Dutch, standard deviations in parentheses.

Prosodic measures	**Hyper-scores**	**18 months (participated at 18 months, *N* = 43)**	**18 months (participated longitudinally, *N* = 27)**	**24 months (participated longitudinally, *N* = 27)**
Word articulation rate	General	0.93 (0.12)	0.90 (0.13)	0.90 (0.11)
	Unfamiliar	0.93 (0.16)	0.93 (0.16)	0.86 (0.11)
Word mean F0	General	1.06 (0.17)	1.09 (0.17)	1.17 (0.17)
	Unfamiliar	1.05 (0.20)	1.07 (0.18)	1.13 (0.15)
Word F0 range	General	1.12 (0.35)	1.15 (0.37)	1.26 (0.35)
	Unfamiliar	1.10 (0.48)	1.14 (0.54)	1.40 (0.51)
Utterance articulation rate	General	0.93 (0.12)	0.92 (0.13)	0.93 (0.08)
	Unfamiliar	0.91 (0.16)	0.91 (0.16)	0.91 (0.10)
Utterance mean F0	General	1.17 (0.13)	1.18 (0.14)	1.12 (0.14)
	Unfamiliar	1.15 (0.15)	1.15 (0.17)	1.09 (0.15)
Utterance F0 range	General	1.06 (0.25)	1.04 (0.23)	1.02 (0.24)
	Unfamiliar	1.09 (0.32)	1.12 (0.33)	1.04 (0.28)

Table 5.2. Concurrent correlations among prosodic hyper-scores and children's vocabulary size in Dutch.

Predictors	**Hyper-scores**	**Concurrent Correlations**			
		18 months ($N = 43$)		**24 months ($N = 27$)**	
		Receptive	**Expressive**	**Receptive**	**Expressive**
Word articulation rate	General	-0.07	0.16	-0.04	0.17
	Unfamiliar	-0.06	0.11	-0.09	0.08
Word mean F0	General	-0.24	0.05	-0.02	-0.07
	Unfamiliar	-0.23*	0.05	-0.22	-0.14
Word F0 range	General	0.02	0.17	-0.07	0.07
	Unfamiliar	-0.02	-0.01	-0.11	0.13
Utterance articulation rate	General	0.01	0.09	-0.14	-0.03
	Unfamiliar	-0.15*	0.04	-0.25*	-0.23
Utterance mean F0	General	0.04	0.01	0.02	0.03
	Unfamiliar	-0.02	-0.04	-0.13*	0.00
Utterance F0 range	General	-0.25	-0.24	0.30	0.33
	Unfamiliar	-0.23*	-0.36*	0.38*	0.53*
Vocabulary size	Receptive	1.00	0.53	1.00	0.81
	Expressive	0.53	1.00	0.81	1.00

Note: $* p < 0.05$; $** p < 0.01$; $*** p < 0.001$

Table 5.3. Longitudinal correlations among prosodic hyper-scores at 18 months and children's vocabulary size at 24 months in Dutch.

Predictors (18m)	**Hyper-scores**	**Correlations with vocabulary at 24m (N = 27)**	
		Receptive	**Expressive**
Prosodic hyper-scores			
Word articulation rate	General	-0.05	-0.06
	Unfamiliar	-0.12	-0.08
Word mean F0	General	-0.12	-0.02
	Unfamiliar	-0.22	-0.07
Word F0 range	General	-0.07	-0.04
	Unfamiliar	-0.09	-0.05
Utterance articulation rate	General	-0.05	0.00
	Unfamiliar	-0.18	-0.09
Utterance mean F0	General	0.11	0.04
	Unfamiliar	0.11	0.04
Utterance F0 range	General	0.06	-0.12
	Unfamiliar	0.13	-0.05
Vocabulary scores			
Receptive 18m		0.80***	0.69***
Expressive 18m		0.40*	0.55*
Receptive 24m			0.81***
Expressive 24m		0.81***	

*Note: * p < 0.05; ** p < 0.01; *** p < 0.001*

5.3.1.1 Concurrent correlations at 18 months: the relationship between the prosodic input to 18-month-old children and children's vocabulary size at 18 months

General hyper-scores

One outlier was capped at the 95th percentile. **Table 5.2** shows simple correlations (Pearson correlation coefficients) between hyper-scores and children's vocabulary size at 18 months and 24 months. In the multiple regression analysis, the final model showed no significant predictors ($R^2 = 0.06$, $F(1, 41) = 2.64$, $p = 0.112$) among the general hyper-scores at 18 months for receptive vocabulary at 18 months. Similar results were obtained when the outcome variable was expressive vocabulary: there were no significant predictors in the final model ($R^2 = 0.11$, $F(2, 40) = 2.54$, $p = 0.091$). These results suggest that there were no concurrent correlations between the general hyper-scores at 18 months and children's receptive or expressive vocabulary sizes at the same age.

Unfamiliar hyper-scores

Six outliers were capped at the 95th percentile. For receptive vocabulary, the final model showed that there were three significant predictors in the final model: utterance articulation rate unfamiliar hyper-score (r = -0.146, β = -382.09, SE = 161.88, t = -2.36, p = 0.024), word mean F0 unfamiliar hyper-score (r = -0.23, β = -164.21, SE = 79.41, t = -2.07, p = 0.046), and utterance F0 range unfamiliar hyper-score (r = -0.23, β = -149.30, SE = 58.61, t = -2.54, p = 0.015). Together these predictors explained 22.7% of the variance (R^2 = 0.227, F(4, 38) = 2.79, p = 0.040). It should be noted that a larger articulation rate indicates faster speech, and hyper-scores were calculated by dividing IDS articulation rate by ADS articulation rate. Thus, a smaller articulation rate hyper-score indicates a slower articulation rate in IDS compared to ADS. As such, these results suggest that mothers who have a slower utterance articulation rate, a lower word mean pitch, and a smaller utterance pitch range when introducing unfamiliar words in IDS tend to have children with a larger receptive vocabulary at 18 months.

With regard to expressive vocabulary, the final model showed that utterance F0 range unfamiliar hyper-score was negatively correlated with vocabulary size (β = -149.30, SE = 58.61, t = -2.547, p = 0.015). This predictor explained 12.7% of the variance (R^2 = 0.127, F(1, 41) = 5.938, p = 0.019). These results suggest that a smaller utterance pitch range in IDS significantly predicts a child's larger expressive vocabulary at 18 months.

5.3.1.2 Concurrent correlations: the relationship between the hyper-scores at 24 months and children's vocabulary sizes at 24 months

General hyper-scores

The results of general hyper-scores for 24 months were similar to the results for 18 months. The final model showed that there was no significant correlation between the general IDS prosody and children's receptive vocabulary (R^2 = 0.16, F(2, 24) = 2.35, p = 0.117). Also, there was no significant correlation between the general hyper-scores at 24 months and the expressive vocabulary at 24 months (R^2 = 0.11, F(2, 40) = 2.54, p = 0.090).

Unfamiliar hyper-scores

One outlier was capped at the 95th percentile. In the final model, three unfamiliar hyper-scores significantly predicted receptive vocabulary at 24 months, and the total R^2 was 0.36 (F(3, 21) = 3.92, p = 0.023). The three significant predictors were: utterance articulation rate unfamiliar hyper-score (r = -0.25, β = -390.54, SE = 175.98, t = -2.22, p = 0.038), utterance mean F0 unfamiliar hyper-score (r = -0.13, β = -281.18, *SE* = 120.59, *t* = -2.31, *p* = 0.031), and utterance F0 range unfamiliar hyper-score (*r* = 0.38, *β* = 159.15, *SE* = 58.39, *t* = 2.73, *p* = 0.013). These results suggest that a slower articulation rate, a lower utterance mean F0, and a larger utterance F0 range specifically when introducing unfamiliar words at 24 months were correlated with children's receptive vocabulary size at 24 months.

As for expressive vocabulary, the final model showed a significant correlation between the unfamiliar hyper-scores and children's expressive vocabulary at 24 months (R^2 = 0.41, $F(3, 21)$ = 4.95, *p* = 0.009). In the model, utterance F0 range unfamiliar hyper-score was positively correlated with vocabulary size (*r* = 0.53, *β* = 242.11, *SE* = 68.25, *t* = 3.46, *p* = 0.002). These results suggest that a larger utterance F0 range when introducing unfamiliar words to 24-month-old children was correlated with children's expressive vocabulary size at 24 months.

5.3.1.3 Longitudinal correlations: the relationship between the hyper-score at 18 and children's vocabulary sizes at 24 months

General hyper-scores

For receptive vocabulary, two outliers were capped at the 95th percentile. **Table 5.3** shows simple correlations (Pearson correlation coefficients) between hyper-scores at 18 months and children's vocabulary size at 24 months. The final model showed that only receptive vocabulary at 18 months significantly predicted children's receptive vocabulary size at 24 months ($F(1, 25)$ = 44.86, $p < 0.001$), with an R^2 of 0.64. None of the prosodic hyper-scores remained significant in the final model. Similarly, only expressive vocabulary at 18 months significantly predicted expressive vocabulary at 24 months (R^2 = 0.31, $F(1, 25)$ = 11.08, *p* = 0.003).

Unfamiliar hyper-scores

Four outliers were capped at the 95th percentile. The results showed two significant predictors for children's receptive vocabulary at 24 months in the final model: utterance F0 range unfamiliar hyper-score (*r* = 0.13, *β* = 102.64, *SE* = 38.57, *t* = 2.66, *p* = 0.014) and children's receptive vocabulary at 18 months (*r* = 0.80, *β* = 0.76, *SE* = 0.10, *t* = 7.84, $p < 0.001$). These two predictors accounted for 72% of children's vocabulary at 24 months (R^2 = 0.72, $F(2, 24)$ = 31.43, $p < 0.001$). Compared to a model with only receptive vocabulary at 18 months as a predictor (R^2 = 0.64, $F(1, 25)$ = 44.86, $p < 0.001$), adding F0 range unfamiliar hyper-score as a predictor improved the model by explaining 8% more of the variance. These results indicate that children who had a larger receptive vocabulary size at 18 months and whose mothers used a larger utterance F0 range when introducing unfamiliar words also had a larger receptive vocabulary at 24 months. When the predictor was expressive vocabulary, the results showed that only children's expressive vocabulary (r = 0.55, β = 1.082, SE = 0.33, t = 3.33, p = 0.003) at 18 months significantly predicted children's expressive vocabulary at 24 months (R^2 = 0.30, F(1, 25) = 11.08, p = 0.003).

5.3.2 Experiment 2: Mandarin Chinese

As we adopted a cross-sectional design in the Mandarin Chinese experiment, we included both age groups (18 months and 24 months) in the model and added Age as a predictor.[25] Also, as M-CDI only evaluates expressive vocabulary, we only examined whether mothers' hyper-scores were correlated with children's expressive vocabulary. **Table 5.4** shows the mean general hyper-scores and mean unfamiliar hyper-scores in Mandarin Chinese.

Table 5.4. Mean general hyper-scores and mean unfamiliar hyper-scores in Mandarin Chinese, standard deviations in parentheses.

Prosodic measures	Hyper-scores	18months	24 months
Word articulation rate	General	0.98 (0.23)	1.02 (0.20)
	Unfamiliar	1.03 (0.29)	0.96 (0.19)
Word mean F0	General	1.13 (0.14)	1.05 (0.11)
	Unfamiliar	1.15 (0.15)	1.06 (0.15)
Word F0 range	General	1.21 (0.33)	1.04 (0.29)
	Unfamiliar	1.17 (0.39)	1.18 (0.38)
Utterance articulation rate	General	1.01 (0.19)	1.00 (0.14)
	Unfamiliar	1.05 (0.26)	0.96 (0.18)
Utterance mean F0	General	1.12 (0.11)	1.03 (0.10)
	Unfamiliar	1.14 (0.12)	1.04 (0.12)
Utterance F0 range	General	1.12 (0.22)	1.01 (0.28)
	Unfamiliar	1.11 (0.42)	1.08 (0.36)

[25] An example of the R code is: CDI_Produced_raw ~ utterance_AR_genhyper + word_AR_genhyper + utterance_mean_genhyper + word_mean_genhyper + word_range_genhyper + utterance_range_genhyper + Age

Table 5.5. Correlations among prosodic hyper-scores and children's vocabulary size for Mandarin Chinese.

Predictors	**Hyper-scores**	**Concurrent Correlations** **18 months Expressive Vocabulary ($N = 17$)**	**24 months Expressive Vocabulary ($N = 16$)**
Word articulation rate	General	0.12	-0.01
	Unfamiliar	0.39	-0.05
Word mean F0	General	-0.09	-0.20
	Unfamiliar	-0.18	-0.32
Word F0 range	General	-0.22	-0.25
	Unfamiliar	-0.09	-0.29
Utterance articulation rate	General	0.36	0.15
	Unfamiliar	0.33	0.01
Utterance mean F0	General	-0.18	-0.38
	Unfamiliar	-0.25	-0.46
Utterance F0 range	General	-0.31	-0.44
	Unfamiliar	-0.02	-0.44

General hyper-scores

Observations from four 18-month-old children and three 24-month-old children were excluded from analyses due to missing vocabulary information, resulting in a total of 33 participants (18 months old: $N = 17$; 24 months old: $N = 16$). **Table 5.5** shows simple correlations (Pearson correlation coefficients) between hyper-scores and children's vocabulary size for Mandarin Chinese. The results from the multiple regression analyses showed that the only significant predictor was Age ($\beta = 316.71$, $SE = 65.19$, $t = 4.86$, $F(4, 28) = 8.41$, $p < 0.001$), with an R^2 of 0.55, indicating that 24-month-old children have a significantly larger expressive vocabulary than 18-month-old children (18 months: mean = 237.12, $SD = 211.17$; 24 months: mean = 542.19, $SD = 151.61$). However, none of the prosodic hyper-scores significantly predicted children's expressive vocabulary.

Unfamiliar hyper-scores

Observations from an additional of three 18-month-old children were excluded from the unfamiliar hyper-score analyses as their mothers did not report any words to be unfamiliar. As a result, the remaining usable sample size was 30 (18-month-old $N = 17$; 24-month-old $N = 13$). The results for unfamiliar hyper-scores were consistent with general hyper-scores. The only significant predictor for children's expressive vocabulary was Age ($\beta = 239.96$, $SE = 72.23$, $t = 3.32$, $F(2, 27) = 11.83$, $p = 0.003$), indicating that 24-month-old children have a larger expressive vocabulary than 18-

month-old children. In the final model, the correlation between utterance mean F0 unfamiliar hyper-score and expressive vocabulary size approached significance (β = -494.07, SE = 266.81, t = -1.852, $F(2, 27)$ = 11.83, p = 0.075).

5.4 Discussion and conclusions

To better understand the role of IDS prosody in children's lexical development, we collected IDS and ADS speech data when mothers addressed their 18- or 24-month-old child vs. an adult in Dutch and Mandarin Chinese. We also collected children's vocabulary information at both ages. The main goal of this study is to explore whether either the generally exaggerated prosody of IDS or IDS prosody specifically in word-learning contexts could predict children's vocabulary size. Thus, we did two sets of multiple regression analyses between the prosodic measurements and children's vocabulary size for each language. First, we examined correlations between the general prosodic exaggeration in IDS and children's vocabulary size. Second, we examined correlations between the prosodic exaggerations specifically when introducing unfamiliar words in IDS and children's vocabulary size. For both Dutch and Mandarin Chinese, we investigated the concurrent correlations between IDS prosody and children's vocabulary size. In addition, we examined the longitudinal correlations for the Dutch longitudinal data in order to show whether prosodic input at 18 months may predict children's vocabulary growth between 18 and 24 months.

First, we asked whether the general prosodic exaggeration in IDS (indicated by general hyper-scores) predicts children's vocabulary size. For Dutch, our results showed no significant correlations between the general hyper-scores and children's vocabulary size in either the concurrent correlations or the longitudinal correlations. Similarly, results for Mandarin Chinese showed that none of the general hyper-scores could significantly predict children's vocabulary size. In their study using a similar hyper-score measure, Kalashnikova and Burnham (2018) found no significant correlations between the mean F0 general hyper-scores addressed to children of 7, 9, 11, 15, and 19 months and children's expressive vocabulary at 15 or 19 months. Here, we extended their analyses to Dutch and Mandarin Chinese and to IDS addressed to two-year-old children and we also found no significant correlations. Previous studies have shown that slower articulation rate and F0 range measured on the global prosody of IDS (in English) are positively correlated with children's larger vocabulary size (Porritt et al., 2014; Raneri, 2015). These studies used raw speaking rate or F0 measures as predictors instead of hyper-scores. It is possible that the inconsistent results are related to the choice of predictors. Also, as the input measures in these studies were obtained from IDS addressed to younger children (3 to 14 months and 7 months, respectively), it is possible that the associations between the prosodic input and children's vocabulary size that were found in these studies are no longer robust at the global level when children are 18 to 24 months of age. Moreover, the null results may be simply due to small sample size in the Dutch and Mandarin Chinese group. Thus, further research should be undertaken to investigate whether the generally exaggerated prosody of IDS is correlated to children's vocabulary size by including various predictors, more age ranges, and larger sample sizes than we were able to.

Although the generally exaggerated prosody of IDS did not seem to be linked to children's vocabulary size, we further examined prosodic exaggeration specifically in word-learning contexts in which mothers introduce unfamiliar words to children. As IDS in word-learning contexts is immediately relevant for children to learn words, we predicted that prosodic exaggeration in this specific context (indicated by the unfamiliar hyper-scores) might be correlated with children's vocabulary size.

Our Dutch results revealed concurrent correlations between unfamiliar hyper-scores and children's vocabulary size at 18 months and 24 months, as well as longitudinal correlations between unfamiliar hyper-scores at 18 months and children's vocabulary size at 24 months. We have three main findings. First, when mothers introduce unfamiliar words to 18-month-old children, a slower utterance articulation rate, a lower word mean F0, and a smaller utterance F0 range together predicted children's larger receptive vocabulary at the same age. Also, a smaller F0 range significantly predicts larger expressive vocabulary at 18 months. Second, at 24 months, a slower utterance articulation rate and a lower utterance mean F0 in IDS in word-learning contexts predicted children's larger receptive vocabulary at 24 months. Regarding F0 range, opposite to the results for 18 months, a larger F0 range significantly predicted larger expressive vocabulary at 24 months. Third, when controlling for the vocabulary size at 18 months, a larger F0 range of utterances with unfamiliar words at 18 months was associated with children's receptive vocabulary 24 months. In other words, a larger utterance F0 range unfamiliar hyper-score at 18 months significantly predicts a larger vocabulary growth from 18 to 24 months.

One question that arises from these results is why these specific prosodic cues in word-learning contexts are correlated with children's vocabulary size. The results of Dutch utterance articulation rate unfamiliar hyper-score and word/utterance F0 unfamiliar hyper-score may be explained by how Dutch mothers modify prosody when they introduce unfamiliar words to children. In Chapter 2 and Chapter 3 of this dissertation, using the same speech data as in this chapter, we found that at the group level, Dutch mothers specifically had a slower utterance articulation rate and a lower utterance/word mean pitch when introducing unfamiliar words. As such, the specific prosodic cues that Dutch mothers used when introducing unfamiliar words (at group level) were also correlated with vocabulary size concurrently.

Regarding F0 range, our results showed significant correlations between F0 range unfamiliar hyper-scores and children's vocabulary size. However, the directions of the correlations for 18- and 24-month-old children differ. In particular, the utterance F0 range unfamiliar hyper-score was negatively correlated with children's receptive vocabulary at 18 months, but at 24 months, it was positively correlated with children's receptive vocabulary. Moreover, a larger F0 range hyper-score at 18 months significantly predicted children's vocabulary growth from 18 to 24 months. As we tested the same mother-child dyads longitudinally, the different results at different ages may not be attributed to sampling. One possible explanation is that the function of F0 range in word-learning contexts was different at different ages. Many studies have shown that the general prosody of IDS changes with age (e.g., Kitamura et al., 2002; Liu, Tsao, & Kuhl, 2009). However, little is known about whether the functions of IDS pitch in children's linguistic development also change with children's age. Our results provide the first evidence that the directions of the correlations between IDS prosody and children's vocabulary size change from 18 months to 24 months.

As in most studies investigating the role of input quantity and input quality in children's vocabulary development, we could interpret our results as supporting the potential facilitative effects of IDS prosody on children's lexical development. However, an alternative account for the significant correlations between input and children's vocabulary could be that mothers adapt their IDS prosody when introducing unfamiliar words according to children's vocabulary knowledge. In this sense, our Dutch results on the concurrent correlations at 18 months might suggest that mothers speak slower, have a lower word pitch, and reduce F0 range when they introduce unfamiliar words to children with a relatively larger receptive vocabulary.

Thus, we have to be cautious with interpreting our results as evidence for the facilitative effects of IDS prosody on children's vocabulary learning. Whether these specific prosodic cues are indeed facilitating children's lexical development, as well as the potential mechanisms by which these prosodic cues facilitate lexical development, requires further investigation. One possibility is that the significant predictors in our Dutch findings, such as slowing down, lowering pitch, and enlarging pitch range specifically when introducing unfamiliar words, together show Dutch mothers' "didactic" intent. Children become sensitive to such speech patterns in word-learning contexts and gradually rely on these prosodic cues to learn new words. As a result, children who are more sensitive to such didactic intent also become better word-learners. Previous studies have examined how prosody is used in IDS for different communicative intents in mother-child interactions in the first year of life (Kitamura & Burnham, 2003). Their results suggest that a higher mean F0 is associated with showing positive affect, expressing affection, encouraging attention, and comforting, while a higher F0 range is related to encouraging attention and directing behavior. They also found that mothers' communicative intent changes with age. As their study focused on the first year of life, the function of pitch in their findings might not be generalized to 18- to 24-month-old children. Even though no study has investigated the didactic intent in mother-child interactions, it is possible that Dutch mothers slow down, lower pitch, and reduce or enlarge pitch range (depending on the age of children) for didactic purposes. Future studies may further investigate whether the prosodic modifications in word-learning contexts indeed support a didactic intent and whether children could benefit from such communicative style.

Our study is the first to explore the relationship between IDS prosody in word-learning contexts and children's vocabulary size. Our results suggest that the general exaggeration of prosody in IDS does not seem to be associated with children's vocabulary size in either Dutch or Mandarin Chinese. Crucially, it is instead IDS prosody specifically when mothers introduce unfamiliar words to children that is predictive of children's vocabulary size. These findings contribute to our understanding of the function of IDS prosody in children's lexical development. Even though extensive studies have been devoted to examining the prosodic characteristics of IDS in comparison to ADS, less is known on the function of IDS prosody in language acquisition and specifically in lexical development. Previous studies have suggested that IDS prosody may only maintain children's attention or facilitates mother-child communication and may not be directly relevant for facilitating language acquisition. For example, the general pitch exaggeration in Australian English IDS did not correlate with children's vocabulary size in Kalashnikova and Burnham (2018). In a meta-analysis drawing from existing studies, Spinelli, Fasolo, and Mesman (2017) suggest that the generally exaggerated pitch of IDS has more attentional functions and less linguistic functions, especially for younger infants before one year of age. Our results support the view that the general exaggeration of prosody in IDS may not be related to vocabulary development. However, our findings on Dutch IDS point to an intriguing possibility that IDS prosody in word-learning contexts is predictive of children's lexical development.

Our study also extends our understanding of the role of IDS in language acquisition to languages other than English. We selected two languages that have typologically different prosody. Dutch, a stress-timed and non-tonal language and Mandarin Chinese, a syllable-timed and tonal language. The significant correlations between Dutch utterance articulation unfamiliar hyper-scores as well as utterance/word pitch unfamiliar hyper-scores and children's vocabulary size support our hypotheses that unfamiliar hyper-scores, but not general hyper-scores, are associated with children's vocabulary size. As illustrated above, these results might be explained by the prosodic modifications Dutch mothers use when introducing unfamiliar words. Our Mandarin Chinese results, however, showed no significant correlations between the unfamiliar hyper-scores and children's expressive vocabulary size. There was only a tendency of significant negative correlation between utterance mean F0 unfamiliar hyper-score and expressive vocabulary size. A possible explanation is that the Mandarin Chinese version of CDI only included expressive vocabulary. Our Dutch results showed that IDS prosody may be more relevant for receptive vocabulary than expressive vocabulary. Thus, it is possible that at these ages, Mandarin Chinese mothers' input is more related to receptive vocabulary. However, so far no instrument has been available to evaluate receptive vocabulary in Mandarin Chinese. Also, we had a relatively small sample size for the Mandarin Chinese experiment. We had 33 participants in total in the two age groups after excluding unusable data. In sum, future research may further investigate whether Mandarin Chinese IDS prosody is correlated with children's receptive vocabulary with a larger sample size.

Taken together, our findings have revealed a link between IDS prosody and children's lexical development, at least for the Dutch dyads under investigation. Specifically, instead of generally exaggerated IDS prosody, it is only the IDS prosody specific to word-learning contexts that is associated with children's lexical development and can explain the variance of children's vocabulary knowledge at 18 months and 24 months. Our findings suggest that the prosody of IDS in word learning contexts may serve specific linguistic purposes.

Chapter 6

Does infant-directed speech facilitate word-to-object mapping for Dutch two-year-old children? Exploring the link between IDS prosody and children's online word learning performances

Abstract

Prototypical infant-directed speech (IDS) is characterized by exaggerated prosody compared to adult-directed speech (ADS). It has long been claimed that prototypical IDS facilitates word learning compared to ADS. Specifically with respect to word-to-object mapping, a crucial skill in word learning, two studies have shown that prototypical IDS facilitates online word-to-object mapping for American English children. However, it remains unknown whether similar effects occur across different languages. Moreover, no study has investigated whether there are correlations between the prosody of IDS and children's online word-to-object mapping performances from ADS and IDS input. This chapter has two research questions. First, does prototypical IDS facilitate word-to-object mapping in 24-month-old Dutch children? Second, are there correlations between the prosody of IDS and children's online word-to-object mapping performances in ADS and IDS? To answer these questions, we collected Dutch IDS speech data in a semi-spontaneous storybook-reading task in which mothers introduced both familiar and unfamiliar words to their 18- and 24-month-old children. We then tested the same children's online word-to-object mapping performances in both ADS and IDS. Results suggest that Dutch 24-month-old children could reliably learn novel words from both ADS and IDS, although IDS has a slight facilitative effect on word learning compared to ADS. There were no significant correlations between the prosody of IDS and children's online word learning performances. These findings suggest that even when children can reliably learn novel words from ADS, the facilitative effects of prototypical IDS on children's online word learning might still exist. However, whether these findings hold across languages or age groups requires further investigation.

6.1 Introduction

Prototypical infant-directed speech (IDS), which is characterized by exaggerated prosody compared to adult-directed speech (ADS), has long been claimed to support children's word learning (e.g., Ferguson, 1964; Fernald et al., 1989). To learn the meaning of a word, children must be able to map a word label from auditory input to a referent, either an object or an action. Despite the widely accepted view that prototypical IDS facilitates children's word learning compared to ADS, only a few studies have provided empirical evidence by directly comparing children's online word learning in ADS and IDS (Mani & Pätzold, 2016; Singh, Nestor, Parikh, & Yull, 2009; Thiessen, Hill, & Saffran, 2005). With regard to word-to-object mapping, the focus of the current study, only two studies have shown that IDS facilitates word-to-object mapping compared to ADS in American English children, at least until they reach 21 months (Graf Estes & Hurley, 2013; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011, to review later). Given that the degree of prosodic exaggeration in IDS might vary across languages (Fernald et al., 1989), it remains unknown whether these results can be generalized to other languages. The first goal of the current study is to extend this research to Dutch-learning children by examining whether prototypical Dutch IDS facilitates 24-month-old children's word-to-object mapping.

Studies that compared children's word learning in ADS and IDS invariably evaluate children's online word learning performances at the group level, ignoring individual differences among children as well as the sources of the individual differences. In particular, such studies have focused on children's word learning independently without taking into account the role of language input. There is robust evidence to show that the quantity and quality of input are correlated with children's language outcomes such as their vocabulary size and online word processing ability (e.g., Hart & Risley, 1995; Hoff, 2006; Hoff & Naigles, 2002; Mahr & Edwards, 2018; Newman, Rowe, & Bernstein Ratner, 2015). As a result, it is possible that a child's word learning performances in ADS and IDS could also be correlated with the IDS he or she is exposed to. As the major difference between ADS and prototypical IDS is its prosody, our second goal is to explore the link between individual IDS prosody and children's online word learning performances in both ADS and IDS. Specifically, we ask whether children's word learning performances in ADS and IDS are associated with the degree of prosodic exaggeration in their mother's IDS.

There are two interpretations of the term "IDS" in the literature. First, IDS is often understood as "prototypical IDS," characterized by exaggerated prosody compared to ADS (Cristia, 2013; Soderstrom, 2007). Second, IDS is sometimes used to refer to *any* speech that is directed to a child, showing great variation and not necessarily featuring prosody exaggerated to the same degree across languages and individuals (e.g., Fernald et al., 1989; Han, de Jong, & Kager, 2018b; Kalashnikova & Burnham, 2018; Narayan & McDermott, 2016; see also Chapters 2 and 3 in this dissertation). Here, as in Chapter 5, the term "IDS" can have either of these two interpretations depending on the context; we will use "prototypical IDS" when we highlight its prosodic differences from ADS to avoid ambiguity.

6.1.1 The facilitative effects of prototypical IDS on word learning

Many studies have compared the characteristics of prototypical IDS and ADS and the findings suggest that the major difference between ADS and IDS is its prosody. Specifically, prototypical IDS is characterized by a higher mean pitch, a larger pitch range, and a slower speaking rate compared to ADS (Cristia, 2013). Such an exaggerated speech register enhances infants' attention and conveys positive affect (Cooper & Aslin, 1990; Trainor & Desjardins, 2002). The attentional and affective functions of IDS may benefit mother-child interaction and lead to better linguistic outcomes in children. Furthermore, many researchers have suggested that IDS has linguistic functions beyond its attentional and affective effects. First, vowel hyperarticulation has been found in IDS in many languages, which may enhance children's phonetic perception (Kuhl et al., 1997; Liu, Tsao, & Kuhl, 2009; but see Benders, 2013; Englund, 2017; and Xu Rattanasone, Burnham, & Reilly, 2013 for opposing evidence). Second, some evidence suggests that IDS prosody highlights contextually new words compared to contextually given words (Bortfeld & Morgan, 2010; Fernald & Mazzie, 1991; Fisher & Tokura, 1995) and helps to distinguish words that are familiar vs. unfamiliar to children (Han, de Jong, & Kager, 2018b; Chapters 2 and 3 in this dissertation), which may consequently facilitate children's word learning.

Despite the evidence suggesting that some properties of prototypical IDS may support word learning, only a few studies have directly compared children's word learning performances under ADS and prototypical IDS conditions experimentally. In order to learn words, children need to segment words from continuous speech, recognize familiar words in speech, and finally they must associate a novel label to a novel object. Correspondingly, studies have examined whether speech with properties of prototypical IDS facilitates word segmentation, word recognition, as well as word-to-object mapping compared to speech without IDS-characteristics (ADS-like speech). In these studies, children generally perform better in tasks related to word learning when they hear prototypical IDS compared to ADS. For example, English- and German-learning infants could only segment words from continuous speech when hearing prototypical IDS input but not when hearing ADS input (Thiessen, Hill, &

Saffran, 2005; Mani & Pätzold, 2016). English-learning infants were able to recognize words that they were familiarized with in prototypical IDS, but not when the words were introduced in ADS (Singh et al., 2009). In a broader sense, word segmentation and word recognition are both important word learning skills. The current study, however, focuses on a narrower definition of word learning, namely word-to-object mapping. So far, two studies have shown that prototypical IDS facilitates word-to-object mapping in American-English-learning children.

Ma, Golinkoff, Houston, and Hirsh-Pasek (2011) was the first empirical study to demonstrate the facilitative effects of prototypical IDS on word-to-object mapping. They tested 21- and 27-month-old American-English-learning children in an online word learning task, and compared their learning performances in the ADS and IDS conditions. For the 21-month-old age group, a between-subject design was adopted. Participants were randomly assigned to either the ADS or IDS condition. They used an Intermodal Preferential Looking Paradigm (IPLP) (Hirsh-Pasek & Golinkoff, 1996). Children were presented with two novel word-object pairs one by one during the training phase. The word labels were pseudowords (e.g., "modi") embedded in carrier sentences such as "It's a modi! See the modi. That's the modi. Look what the modi is doing?" Children in both conditions heard the same speech content, while the only difference was the prosody: audio stimuli in the IDS condition had prototypical English IDS properties including a higher pitch and a larger pitch range than the audio stimuli in the ADS condition. In the testing phase, children were presented with two objects: a target object (correct word-object pair) and a distractor (incorrect word-object pair). The dependent measure was the single longest look at the target object and the distractor in the whole testing phase. Successful word learning was indicated by a longer time looking at the target compared to the distractor. The results showed that there was a significant interaction between Condition (ADS/IDS) and Correctness (Target/Distractor) in the mean single longest look in the testing phase. Specifically, children looked significantly longer at the target than at the distractor in the testing phase in the IDS condition, but they did not look significantly longer at the target in the ADS condition. These results indicate that 21-month-old children could reliably learn novel word-to-object mapping when the words were presented in IDS but not in ADS. For 27-month-old children, they only tested the ADS condition. The results showed that children could reliably learn novel words produced in the ADS condition at 27 months of age. Furthermore, the authors examined whether IDS attracts infants' attention in general by comparing the ADS and IDS conditions with respect to the time spent looking at the screen in the training phase. The results showed that the looking time did not differ between the ADS and IDS conditions. However, children looked longer at the screen in the IDS condition compared to the ADS condition in the testing phase. The authors interpret these results as an indication that IDS maintained children's attention compared to ADS in the testing phase precisely because they had learned the word-to-object mapping in IDS but not in ADS.

Graf Estes and Hurley (2013) extended the study by Ma et al. (2011) by including younger children (17 months old). A crucial difference from Ma et al. (2011) was that they tested whether children could learn word-object associations that were presented in isolation. To do so, they adopted a Switch task (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Similar to Ma et al. (2011), they used a between-subject design. Their task consisted of a habituation phase and a testing phase. The audio stimuli were disyllabic pseudowords (e.g., "gabu") and the visual stimuli were two novel objects. In the habituation phase, two word-object pairings were presented. The audio stimuli in the IDS condition had prototypical IDS prosody, including a higher pitch, a larger pitch range, and a longer duration compared to the stimuli in the ADS condition. In the following testing phase, the word-object pairings were either the same or different (switched) compared to the habituation phase. Successful learning of the word-object associations was indicated by a longer looking time at the switched trials in the testing phase compared to the same trials. Their results suggest that 17-month-old children could learn word-object associations in the IDS condition but not in the ADS condition. In addition, the authors examined whether IDS attracts infants' attention in the habituation phase compared to ADS. Similar to the findings in Ma et al. (2011), children did not pay more attention to the IDS stimuli compared to the ADS stimuli during the habituation phase, suggesting that they did not prefer listening to IDS compared to ADS in general, and yet they only learned words in IDS.

Ma et al. (2011) and Graf Estes and Hurley (2013) adopted different word learning paradigms and the audio stimuli differed. The former study used an Intermodal Preferential Looking Paradigm (IPLP), where words were embedded in carrier sentences, while in Graf Estes and Hurley (2013), a Switch task was used and words were presented in isolation. Yet, their findings converge. These two studies suggest that American-English-learning children could only learn novel words from IDS but not (or, at least not reliably) from ADS before at least 21 months of age.

Both Ma et al. (2011) and Graf Estes and Hurley (2013) tested American-English-learning children, while no study to date has provided evidence for the facilitative effects of prototypical IDS on word-to-object mapping in other languages. IDS shows cross-linguistic differences in various aspects, including the degree of prosodic exaggeration and vowel hyperarticulation. First, while the exaggerated prosody of IDS is often considered universal across languages, language-specificity exists (Fernald et al., 1989; Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002). For example, American English IDS is more exaggerated in prosody than British English, French, Italian, German, and Japanese IDS (Fernald et al., 1989). The difference in the degree of prosodic exaggeration in IDS may influence children's online word learning. Previous studies on word segmentation have shown that while American-English-learning children could segment words from continuous speech as early as 7.5 months (Jusczyk & Aslin, 1995), British-English-learning children failed to do so in the same paradigm until 10.5 months old. Interestingly, 10.5-month-old British English children could only succeed in word segmentation tasks when the stimuli was particularly exaggerated in prosody (Floccia et al., 2016). Second, aside from the differences in generally exaggerated prosody, IDS differs across languages in its vowel hyperarticulation. While vowel hyperarticulation has been shown in various languages, including American English, Russian, Swedish, Taiwanese Mandarin (Kuhl et al., 1997; Liu, Tsao, & Kuhl, 2007), vowels in Dutch or Norwegian IDS instead show evidence of *hypoarticulation* (Benders, 2013; Englund, 2017).

Taken together, despite evidence showing that prototypical IDS facilitates word-to-object mapping in American English children under 21 months of age, the conclusions from only two studies both conducted on the same language cannot be automatically generalized to other languages. It is unclear whether such facilitative effects would hold for other languages. Thus, we adapted the experimental set-up of Ma et al. (2011) to conduct a similar experiment with Dutch children and asked whether prototypical IDS facilitates Dutch children's word-to-object mapping.

6.1.2 The relationship between IDS prosody and children's online word learning

Both Ma et al. (2011) and Graf Estes and Hurley (2013) directly compared children's word-to-object mapping in prototypical ADS and IDS. Their main conclusions were drawn from results at the group level. However, individual differences in the children's word learning in ADS and IDS as well as the sources of individual differences were unclear. In particular, it remains unknown whether the prosody of individual mothers' IDS correlates with their children's online word-to-object mapping. Specifically, while extensive research has examined the correlations between input *quantity* and children's language outcomes, the role of input *quality*, especially the prosodic properties of IDS, has received less attention. Also, results from the few studies on the correlations between IDS prosody and children's language outcomes are mixed.

So far, it is well established that the quantity of IDS input a child receives is correlated with his or her language outcomes. Among these studies, the quantity of

IDS is usually measured in total number of words and/or utterances regardless of prosodic characteristics, and vocabulary size is commonly used as a measure of language outcomes. In general, children who receive a larger amount of IDS tend to develop a larger vocabulary size (Hart & Risley, 1995; Hoff, 2006; Hoff & Naigles, 2002; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Mahr & Edwards, 2018). There are now a growing number of studies demonstrating correlations between language input and children's online word processing skill—another important language outcome measure. For example, Hurtado, Marchman, and Fernald (2008) found that the numbers of utterances and word tokens in IDS at 18 months predicted Spanish-speaking children's reaction time in word recognition tasks at 24 months. Similarly, Weisleder and Fernald (2013) found that the amount of speech that was directly addressed to children (i.e., not including overheard speech) at 19 months predicted Spanish-speaking children's word recognition performances and children's vocabulary at 24 months. However, Mahr and Edwards (2018) showed that the number of words in IDS and American-English-speaking children's word recognition speed at 28–39 months was not significantly correlated.

In addition to input quantity, many aspects of input quality, such as lexical richness, syntactic complexity (Hoff & Naigles, 2002), repetitiveness (indicated by type-token ratio, Rowe, & Bernstein Ratner, 2015), and vowel hyperarticulation were all found to be positively correlated with children's vocabulary size (Hartman, Bernstein Ratner, & Newman, 2016; Kalashnikova & Burnham, 2018; Liu, Kuhl, & Tsao, 2003). However, even though IDS is mainly characterized by exaggerated prosody compared to ADS, only a few studies have explored the relationship between the prosody of IDS and children's language outcomes, and the results are inconsistent. For example, IDS F0 range was found to be positively correlated with American English children's vocabulary (Porritt, Zinser, Bachorowski, & Kaplan, 2014). A slow speaking rate in American English IDS at seven months predicted larger expressive vocabulary at two years of age (Raneri, 2015). However, in a recent study, Kalashnikova and Burnham (2018) did not find any correlations between the exaggeration of pitch in Australian English IDS and children's vocabulary size. Thus, it remains unclear whether IDS prosody is indeed correlated with children's vocabulary size and whether these results may hold across languages. While the correlations between IDS prosody and children's vocabulary are uncertain, it is possible that IDS prosody could be linked to other measures of children's language outcomes: for example, online word learning performance. Suttora et al. (2017) explored the relationship between the prosody of IDS and Italian-learning children's online word recognition performances at 15 months. The prosodic measures included mean F0, maximum F0, minimum F0, and speech rate (including pauses) of mothers' speech during mother-child interactions in a laboratory setting. Their results revealed no significant concurrent correlations between the prosody of IDS and children's word recognition accuracy. Similarly, Song, Demuth, and Morgan (2018) did not find any correlations between the prosody of IDS (including mean pitch and pitch range) at 17 months and children's vocabulary size at 19 or 25 months.

In the studies aiming to correlate individual mothers' prosody and their child's vocabulary size, the prosodic measures were all typically global measures of IDS prosody during mother-child interactions (either at home or in laboratory settings) instead of measurements of prosody in specific interactive contexts such as word-learning contexts. Word-learning contexts, in which mothers introduce unfamiliar words to children, provide the most direct input for children to obtain information for word learning. A recent study has compared 24-month-old Dutch children's online word-to-object mapping performances when they learned novel words from their own mother's speech compared to an experimenter's speech (Van Rooijen, Bekkers, & Junge, 2018). In this study, the content of the speech stimuli were the same for the mother and the experimenter, and both speakers were instructed to introduce the novel word labels using prototypical IDS speech. Still, children learned novel words faster from their mothers' speech compared to the experimenter's speech. Thus, it is possible that the prosody of an individual mother's IDS in word-learning contexts is correlated with children's online word learning performances.

To summarize, despite robust evidence of correlations between input quantity and children's language outcomes, the findings are inconsistent regarding the correlation between IDS prosody and children's language outcomes. It is possible that children's online word learning performances in ADS and IDS are correlated with the prosody of IDS they are exposed to. The current study, therefore, explores the relationship between the prosody (global prosody and prosody in word-learning contexts) of IDS that an individual child is exposed to and his or her online word-to-object mapping.

6.1.3 The current study

To summarize, the current study has two goals. Our first goal is to extend Ma et al. (2011) to Dutch, asking whether prototypical IDS facilitates Dutch children's word-to-object mapping compared to ADS. To answer the research question, we replicated the word learning experiment in Ma et al. (2011) and tested Dutch children's word-to-object mapping in both ADS and IDS conditions. Our second goal is to explore the link between individual IDS prosody and children's word learning performances in ADS and IDS by asking whether the degree of prosodic exaggeration in IDS can predict children's performances in word-to-object mapping in ADS and IDS. Specifically, we asked whether the global prosody of IDS and/or the prosody of IDS specific to word-learning contexts are correlated with children's online word learning performances.

This study is part of a larger study on the role of prosodic input in Dutch and Mandarin Chinese (also see Han, de Jong, & Kager, 2018a; Han, de Jong, & Kager, 2018b; Chapters 2, 3, 4, 5 of this dissertation). In the larger study, we examined the prosody of IDS addressed to 18- and 24-month-old children, during which period children's word learning ability gradually improves and their vocabulary size rapidly increases (Bion, Borovsky, & Fernald, 2013; Goldfield & Reznick, 1990). As 18-month-old children are still developing their "fast mapping" skill (the ability to map a novel label and a novel object based on minimal exposure) (Bion, Borovsky, & Fernald, 2013; Nazzi & Bertoncini, 2003), the current study selected the 24-month-old group in order to further investigate the concurrent correlations between individual IDS prosody and children's word-learning performances at this age.

Ma et al. (2011) suggest that learning novel word-to-object mapping in ADS is more difficult than in IDS. Their findings also show that at 21 months of age, children who had a larger vocabulary also performed better in online word learning. As they had a between-subject design and collapsed the ADS and IDS condition in their correlation analysis, it remains unknown whether children's vocabulary size was related to word learning performance in general or specifically related to one experimental condition. By adopting a within-subject design, we are able to explore the following possibilities: (1) If children's online word learning ability is related to their experience with IDS, a mother's IDS prosody may be correlated with their child's better word learning performance in the IDS condition but not the ADS condition; (2) If IDS facilitates children's word learning ability in general, a mother's exaggerated IDS prosody will be correlated with better word learning performances in both ADS and IDS conditions.

The participants in this chapter are a subset of participants who participated in the production experiments reported in Chapter 2 and Chapter 3 in this dissertation.

6.2 Experiment 1: Testing the prosodic exaggeration of IDS

6.2.1 Participants

The participants were twenty-four Dutch-speaking mother-child dyads who had also participated in the production experiment described in Chapter 2 and Chapter 3 as well as a word-learning experiment (Experiment 2 in this chapter, see Section 6.3). The mother-child dyads visited the lab when the children were 24 months (mean age of children = 24;21, age range = 23;30–26;27). The parent-child dyads were recruited from the Utrecht Baby Lab database and were all Dutch native speakers living in the Utrecht area in the Netherlands. All children were typically-developing without report of language impairments or hearing problems. Written informed consent was obtained for all participants.

6.2.2 Materials and procedure

We used a storybook-telling task to elicit semi-spontaneous speech from the mothers. The storybook contained both words that were familiar and words that were unfamiliar to the children. Each mother told the story twice, once to an adult (ADS condition) and once to their child (IDS condition). To examine whether each child was familiar with the target words or not, mothers filled out a word checklist after the experiment to determine whether their child had already known the target words before reading the picture book. This information was coded as Familiarity (Familiar/Unfamiliar) and used in data analyses. Details about the materials and procedures for collecting the speech data are reported in Chapters 2 and 3.

6.2.3 Prosodic measurements

As reported in Chapters 2 and 3, we extracted the following prosodic measures: one temporal measure (articulation rate, measured by syllable/s) and two pitch measures (mean F0 (Hz) and F0 range (semitone (st)). The prosodic measures were calculated for the utterances containing the target words (target utterances) (see Chapter 2 and Chapter 3 in this dissertation for details on annotation and data treatment). The prosodic measurements were then averaged by Condition (ADS/IDS) and Familiarity (Familiar/Unfamiliar) for each mother.

Following Kalashnikova and Burnham (2018), for each mother, we calculated a "general hyper-score" for each prosodic measurement on the utterance level by dividing IDS values by their corresponding ADS values regardless of whether the word was familiar or unfamiliar. This allowed us to control for individual speaking style by using ADS as a baseline. These general hyper-scores indicate how IDS prosody is generally exaggerated compared to ADS regardless of the familiarity of target words. The hyper-scores are: utterance articulation rate general hyper-scores, utterance mean F0 general hyper-scores, and utterance F0 range general hyper-scores.

As in Chapter 5, we also calculated an "unfamiliar hyper-score" for each prosodic measurement. Unfamiliar hyper-scores only included utterances with unfamiliar words and excluded utterances containing familiar words to serve as an indication of how exaggerated IDS is relative to ADS in word-learning contexts in which mothers introduce an unfamiliar word to their child. The hyper-scores are:

utterance articulation unfamiliar hyper-score, utterance mean F0 unfamiliar hyper-score, and utterance F0 range unfamiliar hyper-score.

6.3 Experiment 2: Word learning experiment

In order to evaluate the relationship between the prosody of IDS and children's word learning performance, the participants in Experiment 1 also participated in Experiment 2: a word learning experiment at 24 months. This experiment was adapted from Ma et al. (2011). As in Ma et al. (2011), we used the Intermodal Preferential Looking Paradigm (IPLP) (Hirsh-Pasek & Golinkoff, 1996). A major difference between the current study and Ma et al. (2011) was that they used a between-subject design, but in our study we adopted a within-subject design.

6.3.1 Participants

The participants were the same as those in Experiment 1.

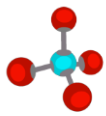
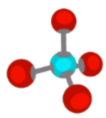
6.3.2 Apparatus

Children were tested in a quiet testing room with a testing area on one side and a control area on the other side. The testing area was separated from the control area with curtains during testing. In the testing area, a 46-inch monitor was about 94 cm from a chair for the parent-child dyads. The height of the screen was at the children's eye level. Parents wore headphones (AKG K109SB with high ambient noise attenuation) playing music so they could not hear any sounds from the experiment. A hidden camera was situated below the screen to record children's visual fixation. Audio stimuli were delivered through a speaker (Tangent EVO-E4) in front of the testing area.

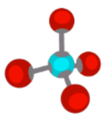
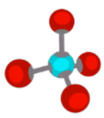
6.3.3 Audio stimuli

Two pairs of phonotactically-legal Dutch disyllabic pseudowords were created as novel labels: Pair 1: "modi" [ˈmodi] and "dofa" [ˈdofa] and Pair 2: "doboe" [ˈdobu] and "pima" [ˈpima]. These two pairs of novel labels were randomly assigned to ADS and IDS across participants. The carrier sentences were adapted from Ma et al. (2011) (see **Table 6.1**).

Table 6.1. An example of the procedure using Picture Pair 1 and Word Pair 1.

Phases	**Left Side**	**Right Side**	**Audio (with English translation)**
Task familiarization			Bal! Kijk naar de bal! Zie je de bal? Dat is de bal. "Ball! Look for the ball! Can you find the ball? That's the ball."[26]
			Boek! Kijk naar het boek! Zie je het boek? Dat is het boek. "Book! Look for the book! Can you find the book? That's the book."
Salience			No audio
Training (4 trials; 20 seconds per trial; Each trial featured a pair of pseudo words and was repeated in ABAB format as such: "modi" – "dofa" –			Kijk daar eens! Het is een modi! Zie je de modi? Dat is de modi. Zie je wat de modi aan het doen is? Nou komt de modi hierheen. Waar gaat de modi naar toe? Waar is de modi? Modi! Daar is de modi! "Look here! It's a modi! See the modi. That's the modi. Look what the modi is doing? Now the modi is

[26] The English scripts were adapted from Ma et al. (2011).

“modi” – “dofa”)			going over here. Where’s the modi going? Where’s the modi? Modi! There’s the modi!”
			Kijk daar eens! Kijk daar eens! Het is een dofa! Zie je de dofa? Dat is de dofa. Zie je wat de dofa aan het doen is? Nou komt de dofa hierheen. Waar gaat de dofa naar toe? Waar is de dofa? Dofa! Daar is de dofa! “Look here! It’s a dofa! See the dofa. That’s the dofa. Look what the dofa is doing? Now the dofa is going over here. Where’s the dofa going? Where’s the dofa? dofa! There’s the dofa!”
Test block (4 trials; 6 seconds per trial; each word was repeated twice)			Modi! Waar is de modi? Kijk naar de modi! Daar is de modi. “Modi! Where’s the modi? Look at the modi! There’s the modi.”
			Dofa! Waar is de dofa? Kijk naar de dofa! Daar is de dofa. “Dofa! Where’s the dofa? Look at the dofa! There’s the dofa”

The audio stimuli were produced by a female Dutch native speaker. This speaker was instructed to produce scripted sentences in natural ADS and IDS prosody. In the ADS condition the speaker was asked to imagine that she was talking to an adult, and in the IDS condition she was asked to imagine that she was talking to an infant. The speaker recorded 5 versions for each speech register. Three Dutch native speakers who were blind to the speech registers of each recording judged the speech register (ADS/IDS) and the naturalness of the recordings and we selected the most natural version of ADS and IDS. The prosody of stimuli was analyzed in Praat. The audio stimuli in the IDS condition had a higher pitch, a larger pitch range, and a slower articulation rate compared to ADS (see **Table 6.2** for the prosodic characteristics of the ADS and IDS stimuli). As in Ma et al. (2011), the length of the recordings differed between the ADS and IDS recordings such that the natural ADS speech samples were generally shorter than IDS speech. We followed Ma et al. (2011) and inserted silent pauses between utterances in the ADS samples to ensure that the total duration of recording was the same between the two conditions samples (see **Table 6.2** for mean pause duration of the ADS and IDS stimuli). The naturalness of the edited speech samples was again confirmed by two Dutch native speakers.

Table 6.2. Prosodic measures of audio stimuli.

Prosodic measures	**Condition**	**Mean (*SD*)**
Mean F0 (Hz)	ADS	217 (10.2)
	IDS	304 (12.0)
F0 range (Hz)	ADS	275 (82.7)
	IDS	440 (24.3)
Articulation rate (syllables/s)	ADS	4.78 (0.20)
	IDS	4.25 (0.32)
Mean pause duration (s)	ADS	1.00 (0.05)
	IDS	0.83 (0.12)

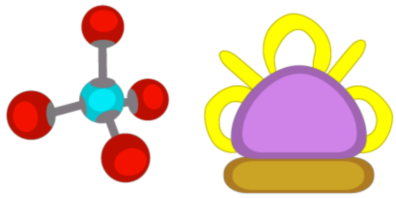
Note: We used linear mixed-effects models to check the prosodic differences in the stimuli. The fixed factor was Condition (ADS/IDS) and Word (Doboe/Dofa/Modi/Pima) was a random factor. The results show significant main effects of Condition on mean F0 (β = 87.837, SE = 5.57, t = 15.78, $p < 0.001$), F0 range (β = 165.38, SE = 31.66, t = 5.224, $p < 0.001$), articulation rate (β = -0.17, SE = 0.43, t = -3.93, p = 0.002), and mean pause duration (β = -0.536, SE = 0.13, t = -4.16, p = 0.002).

6.3.4 Visual stimuli

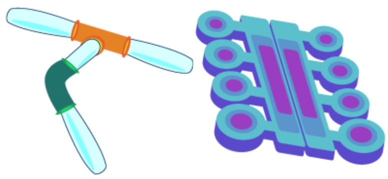
The familiar objects were the same as in Ma et al. (2011): a book and a ball. We designed two pairs of novel objects. The novel objects in each pair are distinctive from each other in both shape and colors. Objects in Pair 1 were adapted from (Liu & Kager, 2018). Objects in Pair 2 were created by the author.

Figure 6.1 Novel objects.

Pair 1:



Pair 2:



6.3.5 Procedure

Participants visited the Utrecht Baby Lab twice on two different days after they reached 24 months (24;00). The average duration between the two visits was 7.6 days (*SD* = 7 days, range = 2–35 days). They were tested on ADS or IDS separately during the two visits. The order of the two conditions was counterbalanced across participants. During the experiment, children sat on their caregiver's (all mothers) lap facing the monitor. Parents were instructed to hold their child without interacting with him/her, pointing to the screen, or talking to him/her. Following the procedure in Ma et al. (2011), each session consisted of five phases: task familiarization, salience, training, and testing phases (**Table 6.1**).[27]

Before each trial, a picture of a baby face appeared in the center of the screen accompanied by a giggling sound of a baby to attract children's eye fixation to the center of the monitor. Which side appeared first and the object-sound pairing were also randomized across the participants. After the visual stimuli were shown, there was a two-second silence before the audio stimuli was played. Each session lasted for about five minutes.

[27] There were two test blocks in the original experiment. Due to an error in the experiment scripts, Test Block 2 of half of the participants could not be used. We therefore excluded Test Block 2 from our final analysis and only drew conclusions from Test Block 1.

6.3.6 Coding and data analysis

Two coders did a pre-screening of the video data to check whether the parent interfered with the child during the testing session. Coding was done with the UiL OTS Video Coding System (De Mooij, 2017). The first coder coded the looking directions (left, center, right, or distracted) frame-by-frame throughout each video. The coder was blind to the condition and trial. Also, as the video was silent, the coder could not hear the audio stimuli. All the coding was then checked by a second coder (the first author) who was also blind to the condition and trial, and the agreement was 98.5%. Any obvious errors were corrected. Looks in which looking directions could not be decided were marked as "unknown" and excluded from analyses.

To evaluate children's word-to-object mapping in ADS and IDS, we had three dependent measures. First, as in Ma et al. (2011), we also calculated the single longest look at the target and to the distractor for each testing trial. Similar to the original study, we did not specify any time window for children's looking behavior but instead determined the single longest looks in the entire trial (6000 ms). In addition to this measure, we had two other dependent measures commonly used for IPLP: proportion of looking time and latency of shift to target from the distractor (Delle Luche, Durrant, Poltrock, & Floccia; 2015; Fernald, Zangl, Portillo, & Marchman, 2008). Previous studies suggest that the classic time window for the naming effect to take place starts from 367 ms after the onset of the naming, and children's looking behavior is at chance level after roughly 2000 ms after the word onset. The choice for this time window in practice, however, tends to differ among studies depending on the task and the age of the children. Thus, instead of calculating the looking behavior of the whole testing trials, we adopted a time window (specifically, 200–2200 ms) that had been used in a previous study with a similar experimental design and on the same age group to test whether mothers' IDS facilitated Dutch 24-month-old children's novel word learning compared to a stranger's voice (Van Rooijen, Bekkers, & Junge, 2018). As in their study, the proportions of looking time to target and to non-target were calculated in a two-second time window which began 200 ms after the onset of the novel labels in each trial. As the audio stimuli in the testing phase started with the novel label, the beginning of the two-second time window was marked at 200 ms after the onset of the audio stimuli. The formula for calculating the proportion of looking time is: Proportion of looking time = (Target)/(Target + Distractor). The latency measure indicates how quickly children shift from the distractor to the target 200 ms after they hear the target words. For this measure, only the trials in which children began by looking at the distractor were included in the analysis. In total, 74 out of 192 trials were included in this latency measure analysis.

We used the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in the R environment (R Core Team, 2018) for analysis. To examine whether children learn novel word-to-object mappings better in IDS than in ADS, we used linear mixed-effects models for all analyses. In the models, we included fixed factors of Condition (ADS/IDS) and Target (Target/Distractor) with Participant as a random factor. We included Condition and Target as random slopes (Barr, Levy, Scheepers, & Tily, 2013).[28] Then, we used the "step" function in the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) to reduce the models by eliminating non-significant factors or interactions. The dependent measures were single longest look (ms), proportions of looking time (%), and latency (ms).

We first detected outliers by visual inspection of scatter plots and capped them at the 5th (for outliers below the lower limit) or the 95th percentile (for outliers above the upper limit). To explore whether there are correlations between IDS prosody and children's word learning performance for ADS or IDS, we first calculated simple correlations (Pearson correlation coefficient) between the hyper-scores (general hyper-scores and unfamiliar hyper-scores) and children's word-learning performance scores using the cor() function in the R Stats Package (R Core Team, 2018). The *p*-values were calculated using the rcorr() function in the Hmisc package (Harrell, 2018). We further conducted a series of multiple regression analyses. The multiple regressions were done in the R environment (R Core Team, 2018) using the lm() function. For each model, we started by including all the predictors[29] and then used the "stepAIC" function in the MASS package (Venables & Ripley, 2002) to reduce the model by selecting variables with a significance level of 5% (direction was set to "backward"). We also checked multicollinearity among the predictor variables. If two predictor variables were highly correlated ($r > 0.8$), we excluded one of the two predictors before we built each model.

With respect to the outcome measures, we averaged the three dependent measures for children's word learning performances from Experiment 1 across trials by condition for each participant. Since we had three dependent measures to evaluate children's word learning performance, we calculated three sets of word-learning performance scores as the outcome measures for the correlation analyses:

1. Differences in mean single longest look between target and distractor. This score was adapted from Ma et al. (2011). As in their study, we

---

28 An example of the R codes is: lmer(lookingtime ~ Condition * Target + (1 + Condition + Target |Participant) + (1 + Condition + Target |Trial))

29 An example of the R codes is: lm(word_learning_performance_score ~ utterance_articulation_rate_hyper + utterance_mean_F0_hyper + utterance_F0_range_hyper)

subtracted the mean single longest look at the distractor from the mean single longest look at the target. As we tested both ADS and IDS conditions within participants, we calculated two scores for ADS and IDS respectively: differences in mean single longest look between target and distractor in ADS and differences in mean single longest look between target and distractor in IDS.

2. Proportions of looking time to the target. For this measure, we used the raw values of the proportions of looking time to the target in ADS and proportions of looking time to the target in IDS.
3. Latency. The latency measure included latency in the ADS condition and latency in the IDS condition.

6.4 Results

6.4.1 Does IDS facilitate word-to-object mapping?

Table 6.3. Means and *SD*s of three dependent measures: single longest look, proportion of looking time, and latency.

Measure	Condition	Target	Mean (*SD*)
Single longest look (ms)	ADS	Target	2509 (843)
		Distractor	1537 (802)
	IDS	Target	2980 (948)
		Distractor	1321 (453)
Proportion of looking time (%)	ADS	Target	58.1 (33.9)
		Distractor	41.9 (33.9)
	IDS	Target	60.3 (32.1)
		Distractor	39.7 (32.1)
Latency (ms)	ADS		836 (429)
	IDS		771 (445)

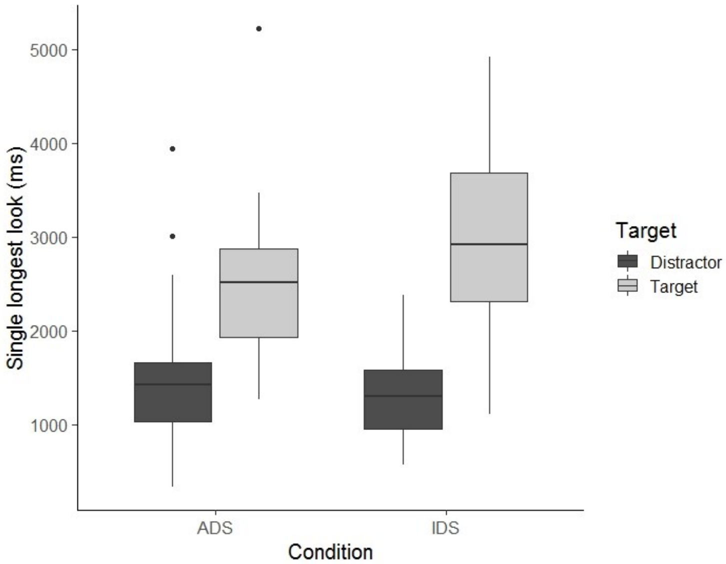
Table 6.3 shows the means and standard deviations of the three dependent measures for Target and Distractor in each condition (ADS/IDS). For the measure "single longest look", we excluded the observations in which children did not look to the target or the distractor (i.e. single longest look = 0 ms). The final model (selected by the step() function) including random factors had very large random variances and non-normally distributed residuals. We therefore decided to carry out two-way repeated measures ANOVA's,[30] after aggregating for each participant by Condition

[30] An example of the R codes is: aov(single_long_mean ~ (Condition*Target) + Error(Participant/(Condition*Target))

and Target with Condition (ADS/IDS) and Target (Target/Distractor) as within-subject variables. **Figure 6.2** shows box plots of children's single longest look to the target or the distractor in ADS and IDS conditions. One outlier (> 3 standard deviations from the mean) was excluded from the analyses.[31] The results showed that there was a significant main effect of Target (Target/Distractor) ($F(1, 22) = 61.49$, $p < 0.001$) and a significant main interaction of Condition (ADS/IDS) and Target ($F(1, 22) = 5.568$, $p = 0.028$); however, the main effect of Condition was not significant ($F(1, 22) = 1.102$, $p = 0.305$). These results indicated that children's single longest look at the target was longer compared to the distractor in ADS, but children's single longest look at the target was specifically longer in IDS compared to ADS. As results did not consistently show a significant interaction between Condition and Target (see footnotes 28), and the effect (if it existed) was rather small, we should be interpreting these results with caution. In sum, the results for single longest look suggest that Dutch 24-month-old children could already learn novel word-to-object mapping in both the ADS and IDS conditions, however IDS might have a small facilitative effect on children's word learning compared to ADS.

[31] When including the observations with single longest look of 0 ms, the results showed that there was a significant main effect of Target (Target/Distractor) ($F(1, 23) = 53.83$, $p < 0.001$). However, there was no significant main effect of Condition (ADS/IDS) ($F(1, 23) = 0.612$, $p = 0.442$), nor was there a significant interaction of Condition and Target ($F(1, 23) = 2.515$, $p = 0.126$). Furthermore, when including the outlier, the results showed that there was a significant main effect of Target (Target/Distractor) ($F(1, 23) = 50.51$, $p < 0.001$) and the interaction of Condition (ADS/IDS) and Target was approaching significance ($F(1, 23) = 5.568$, $p = 0.058$); however, the main effect of Condition was not significant ($F(1, 23) = 1.046$, $p = 0.317$).

Figure 6.2 Box plots of single longest look (ms) to the target and the distractor in ADS and IDS conditions.



We performed two-way repeated measures ANOVA's for proportions of looking time. The results showed that there was only a significant main effect of Target (Target/Distractor) ($F(1, 23) = 12.25$, $p = 0.002$), but there was no significant main effect of Condition (ADS/IDS) ($F(1, 23) = 0.188$, $p = 0.668$), nor was there a significant interaction of Condition and Target ($F(1, 23) = 0.31$, $p = 0.583$). Children looked significantly longer to the target compared to the distractor in both the ADS and IDS conditions, suggesting that children learned the words in both conditions, consistent with the results of single longest look, where we also found a significant main effect of Target. However, for this measure, there was no evidence of facilitative effects of IDS.

We further tested whether children were faster at looking at the target words in the IDS condition compared to the ADS condition. For the dependent variable "latency", no random factors or fixed effects remained in the final model, suggesting that there was no significant main effect of Condition (ADS/IDS), nor for Target (Target/Distractor), nor was there an interaction. As such, there is no evidence to suggest that children looked at the target significantly faster in the IDS condition compared to the ADS condition.

6.4.2 Is IDS prosody correlated with children's word learning performances in ADS and IDS?

Four of the participants were excluded from the correlation analysis as their speech data were not available. As a result, data from 20 participants were included in the final analyses. The mean hyper-scores are shown in **Table 6.4**. As 7 participants did not have latency scores in both ADS and IDS conditions, the remaining sample size (N = 13) was too small for correlation analyses after excluding these participants. Thus, we excluded this measure from the final analyses.

Table 6.4. Mean general hyper-scores and mean unfamiliar hyper-scores in Dutch, standard deviation in parenthesis.

Prosodic measures	**Types of hyper-score**	**Mean hyper-scores (*SD*)**
Utterance articulation rate	General	0.96 (0.07)
	Unfamiliar	0.94 (0.09)
Utterance mean F0	General	1.13 (0.15)
	Unfamiliar	1.07 (0.17)
Utterance F0 range	General	1.07 (0.21)
	Unfamiliar	1.08 (0.30)

Table 6.5. Pearson's correlations between general hyper-scores at 24 months and children's word learning performances (N = 20, p values in parenthesis).

Predictors	**Correlations**			
General hyper-scores	Differences in the single longest look		Proportions of looking time to the target	
	ADS	IDS	ADS	IDS
Articulation rate	-0.06 (p = 0.80)	-0.23 (p = 0.32)	-0.26 (p = 0.26)	0.03 (p = 0.89)
Mean F0	-0.20 (p = 0.39)	-0.16 (p = 0.51)	0.12 (p = 0.60)	-0.32 (p = 0.16)
F0 range	-0.04 (p = 0.85)	0.10 (p = 0.68)	-0.33 (p = 0.16)	-0.05 (p = 0.80)

We first examined whether there were correlations between the general hyper-scores and children's word learning performances in ADS and IDS. As shown in **Table 6.5**, there were no significant correlations between any of the general hyper-scores and the two word learning performance scores. We further examined whether there were correlations between the unfamiliar hyper-scores and children's word learning performances in ADS and IDS. Two participants were excluded as these two participants did not report any unfamiliar words and so no hyper-scores were available. Therefore, 18 participants were included in the analyses of unfamiliar hyper-scores.

As shown in **Table 6.6**, there were no significant correlations between any of the unfamiliar hyper-scores and children's word learning performances in ADS or IDS.

We further performed the planned multiple regressions. As could be expected based on the outcomes of these correlations, none of the models in the multiple regression analyses showed significant relations between the outcome measures and the predictor measures (all *p*'s > .14).

Table 6.6. Pearson's correlations between unfamiliar hyper-scores at 24 months and children's word learning performances ($N = 18$, *p* values in parenthesis).

Predictors	**Correlations**			
Unfamiliar hyper-scores	Differences in the single longest look		Proportions of looking time to the target	
	ADS	IDS	ADS	IDS
Utterance articulation rate	0.03 ($p = 0.89$)	-0.36 ($p = 0.14$)	-0.16 ($p = 0.51$)	-0.02 ($p = 0.94$)
Utterance mean F0	-0.14 ($p = 0.59$)	-0.09 ($p = 0.71$)	0.06 ($p = 0.80$)	-0.22 ($p = 0.37$)
Utterance F0 range	-0.13 ($p = 0.62$)	-0.01 ($p = 0.95$)	-0.26 ($p = 0.30$)	-0.15 ($p = 0.55$)

To summarize, based on our (small) sample size, there were no significant correlations between the hyper-scores, neither general nor unfamiliar, and children's online word learning performances, neither in ADS nor in IDS.

6.5 Discussion and conclusions

The current study had two goals. First, we extended previous studies investigating the facilitative effects of American English IDS in word-to-object mapping to Dutch. Second, we explored whether children's word-to-object mapping performances in ADS and IDS can be predicted by their mothers' IDS prosody.

First, does IDS facilitate Dutch 24-month-old children's word-to-object mapping? We used the same dependent measure "single longest look" as in Ma et al. (2011). Our results show that children looked significantly longer at the target, specifically in the IDS condition, suggesting that IDS indeed facilitates children's word-to-object mapping, however this finding is not robust. Also, our results differ from Ma et al. (2011) regarding children's word learning ability. In the original study, Ma et al. (2011) found that English-learning 21-month-old children could only learn word-to-object mappings when listening to the auditory forms of words presented in prototypical IDS but not in ADS. Only after 27 months could children learn novel words presented in ADS. However, our results showed that Dutch 24-month-old children could already learn novel word-to-object mappings reliably in both ADS and

IDS conditions. It is possible that children at 24 months of age, instead of 27 months of age, could reliably learn novel word-to-object mappings under both conditions.

Even though the results for single longest look suggest that IDS might facilitate word-to-object mapping compared to ADS, our results on the proportion of looking time and latency did not support this view. In particular, results on the proportion of looking time showed that children could learn words in both ADS and IDS, and there was no significant difference in learning performance between the two conditions. Similar results were obtained for the latency measure: children did not switch to the target significantly faster in the IDS condition.

Altogether, we have to interpret our results with caution. For both single longest look and proportion of looking time, our results suggest that Dutch children can reliably learn novel word-to-object mapping at 24 months of age. However, the results for the single longest look and proportion of looking time were inconsistent. Even though several studies have proposed that single longest look is a more sensitive measure compared to total looking time (Schafer & Plunkett, 1998; Bailey & Plunkett, 2002), both measures have been used as valid measures in this paradigm (Fernald et al., 2008; Delle Luche, Durrant, Poltrock, & Floccia, 2015).

Ma et al. (2010) suggest that children can reliably learn novel words in ADS at 27 months. However, as they only tested the ADS condition for this age, it is not clear if there is still a facilitative effect of prototypical IDS at 27 months. Given that we found a small facilitative effect of prototypical IDS on 24-month-old children who could already learn novel words in ADS, the implications for future research are twofold. First, when do children start to reliably learn novel words from ADS? Second, when does the facilitative effect of prototypical IDS on children's word learning decrease? So far, only three studies (including our study) have examined the effect of IDS on children's word learning experimentally, and they examined different languages, used different paradigms, and adopted different measurements. Also, the sample sizes in these studies were all rather small. Thus, these two questions require further empirical evidence from different languages, different ages, and different research methods, or even with a collaborative data collection across laboratories (see Frank et al., 2017 for a collaborative investigation on the listening preference for IDS). Regarding the second question, the facilitative effects of IDS prosody may not be limited to first language acquisition. In a study that investigated whether speech with IDS-like prosody helped L2 learners' word learning, Golinkoff and Alioto (1995) found that English-speaking adults learned Chinese words better when these words were produced in IDS-like speech (exaggerated in prosody) and were placed in utterance-final position, suggesting that properties of IDS (including prosody and word order) may continue to promote second language learning in adults.

Our second research question was whether there were correlations between IDS prosody and children's word learning performances in ADS and in IDS. We had two predictions. First, if children's online word learning ability is shaped by their experience with input, children who are exposed to more exaggerated IDS prosody (indicated by general hyper-scores) and more exaggerated IDS prosody in word-learning contexts (indicated by unfamiliar hyper-scores) will also perform better in word learning in the IDS condition. Second, if IDS facilitates children's word learning ability in general, children who have been exposed to more exaggerated IDS prosody will have better word learning skills, and consequently perform better at word learning in both conditions.

Our results did not show any significant correlations between the general hyper-scores and children's word learning performances or between the unfamiliar hyper-scores and children's word learning performances. Due to the small sample size (N = 20 and N = 18), our correlation analysis was rather exploratory. Thus, it is not possible to draw any conclusions based on our rather small sample size and non-significant results. With these small sample sizes, we were only able to find large correlations. Using the pwr.r.test() function in the pwr package (Champely et al., 2018), we calculated that with a power of 0.8, we would only have been able to find correlations of 0.58 (N = 20) and larger. Therefore, future studies may investigate this issue further by including larger sample sizes than we were able to.

This study revisited the effect of prototypical IDS on children's word learning and explored the relationship between IDS prosody and children's online word learning performance. Our results suggest that 24-month-old Dutch children reliably learn novel word-to-object mapping in the ADS condition, but there is a small facilitative effect of prototypical IDS on children's word learning. Also, there were no significant correlations between IDS prosody and children's online word learning performances in ADS and IDS based on the results from our exploratory correlation analyses. As such, whether prototypical IDS is necessary or beneficial for word learning across languages and throughout early childhood is still an open issue. In order to better understand the role of prototypical IDS in word learning, further investigations on the potential facilitative effects of prototypical IDS need to be carried out on different languages and different age groups.

Chapter 7

General discussion and conclusion

When talking to children, caregivers across the world use infant-directed speech (IDS)—a speaking style which is characterized by exaggerated prosody compared to adult-directed speech (ADS). Despite the long-standing claim that IDS serves to benefit children's word learning (e.g., Snow, 1977), a review of the literature reveals five research gaps regarding the role of prosodic input in word learning. First, no previous study has addressed the prosody of IDS specific to word-learning contexts in which mothers introduce unfamiliar words to children. Second, studies on the prosody of IDS have typically focused on a single language, instead of cross-linguistic investigations of prosodically distinct languages. Third, most studies have focused on IDS addressed to children in their first year of life, while relatively fewer studies have examined IDS during children's "vocabulary spurt" period from 18 to 24 months. Fourth, there is some evidence to suggest that prototypical IDS facilitates online word learning for American English children (Graf Estes & Hurley, 2013; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011), but it remains unknown whether similar effects occur across different languages. Finally, it remains unexplored whether there are correlations between prosodic input in word-learning contexts and children's language outcomes, such as vocabulary size and online word learning performances.

The main goal of this dissertation was to fill the above-mentioned research gaps by investigating the role of prosodic input in children's word learning. To achieve this goal, this dissertation examined the prosody of IDS in two typologically distinct languages: Dutch (a Germanic, stress-timed language with lexical stress) and Mandarin Chinese (a Sinitic, syllable-timed language with lexical tone). As we were particularly interested in the role of prosody in word learning, we focused on the ages 18 months and 24 months, between which children's vocabulary size increases rapidly and their word-to-object mapping skill develops (Bion, Borovsky, & Fernald, 2013; Goldfield & Reznick, 1990). This concluding chapter summarizes the main findings from the five preceding empirical chapters, discusses their general implications, addresses methodological issues, and proposes directions for future research.

7.1 Summary of main findings

7.1.1 Overview of chapters

Chapters 2–6 reported results from three experiments. Experiment 1 examined Dutch IDS prosody in word-learning contexts by eliciting speech data in a semi-spontaneous storybook-telling task. This experiment tested Dutch mothers' IDS prosody longitudinally when their children were 18 months and 24 months. The results of Experiment 1 are reported in Chapters 2, 3, 5, and 6, which focus on different prosodic properties of the speech data. Using similar materials and speech elicitation methods as in Experiment 1, Experiment 2 examined Mandarin Chinese IDS prosody in word-learning contexts. This experiment had a cross-sectional design and tested Mandarin Chinese mothers of 18- and 24-month-old children. The results of Experiment 2 are reported in Chapters 2, 3, 4, and 5, focusing on different aspects of the speech data. Experiment 3 was a word learning experiment adapted from Ma et al. (2011) using an Intermodal Preferential Looking Paradigm (Hirsh-Pasek & Golinkoff, 1996). The results of Experiment 3 are reported in Chapter 6.

Chapters 2 and 3 investigated the prosody of IDS in word-learning contexts in Dutch and Mandarin Chinese, with foci on two aspects of prosody: Chapter 2 examined speaking rate while Chapter 3 focused on the pitch properties of the speech data. These two chapters mainly addressed the following research questions:

> **RQ 1**: What prosodic means do mothers use in IDS in word-learning contexts in which they introduce unfamiliar words to their children?
> **RQ 2**: How does the prosody of IDS in word-learning contexts (compared to ADS) differ between Dutch and Mandarin Chinese?
> **RQ 3**: Does the global prosody of IDS and IDS prosody in word-learning contexts change from 18 to 24 months of age?

Chapter 2 demonstrated that overall, Dutch IDS addressing 18- and 24-month-old children had a slower articulation rate compared to ADS, but there were no significant differences in articulation rate between Mandarin Chinese ADS and IDS in either age group. In word-learning contexts, Dutch mothers slowed down specifically when they introduced unfamiliar words in IDS compared to familiar words, which may potentially facilitate children's word learning. In contrast, there was no evidence to suggest that Mandarin Chinese IDS slows down to highlight unfamiliar words. In both Dutch and Mandarin Chinese, our results showed no evidence of age-related changes in IDS speaking rate addressed to children aged 18 and 24 months.

Chapter 3 showed that the overall pitch exaggeration of IDS differed between Dutch and Mandarin Chinese: Dutch IDS addressed to 18- and 24-month-old children had an overall higher mean pitch than ADS but the differences in pitch level between the two speech registers were less for 24-month-old compared to 18-month-old children. There were no significant differences in pitch range between ADS and IDS for either age group. Mandarin Chinese IDS addressing 18-month-old children had a higher mean pitch and a larger pitch range compared to ADS, but Mandarin Chinese IDS addressed to 24-month-old children was similar to ADS with respect to both mean pitch and pitch range. The prosody in word-learning contexts also differed between Dutch and Mandarin Chinese. In particular, Dutch mothers had a lower mean pitch specifically for unfamiliar words for both 18- and 24-month-old children. Mandarin Chinese mothers, on the other hand, had a higher mean pitch specifically when they introduced unfamiliar words to 18-month-old children, and they had a larger pitch range in IDS specifically when introducing unfamiliar words to 24-month-old children.

As lexical tones are crucial to word meanings in Mandarin Chinese, in addition to the word and utterance prosody investigated in Chapters 2 and 3, **Chapter 4** examined the acoustic features of Mandarin Chinese lexical tones in IDS and addressed the following research questions:

> **RQ 4a**: Are tones in Mandarin Chinese IDS addressed to 18- and 24-month-old children hyperarticulated compared to tones in ADS?
> **RQ 4b**: Does the hyperarticulation of lexical tones in Mandarin Chinese IDS differ when mothers address 18-month-old versus 24-month-old children?

The results of Chapter 4 showed that lexical tones in IDS addressing 18-month-old children had a higher minimum F0, higher maximum F0, and larger pitch range than lexical tones in ADS. Lexical tones in IDS addressing 24-month-old children showed more similarity to ADS tones with respect to pitch height: there were no differences in minimum F0 or maximum F0 between ADS and IDS. However, F0 range was still larger. These findings suggest that lexical tones are hyperarticulated in Mandarin Chinese IDS addressing both 18- and 24-month-old children despite the change in pitch level over time.

After establishing the *nature* of IDS prosody in word-learning contexts, Chapters 5 and 6 further investigated whether mothers' individual IDS prosody is correlated with their children's language outcomes. In both chapters, we calculated prosodic hyper-scores based on the speech data collected in Experiment 1 and Experiment 2 to use as predictors. These two chapters used two different measures of language outcomes: Chapter 5 used children's vocabulary size while Chapter 6 used children's online word learning performances. Chapter 5 included data from both Dutch and Mandarin Chinese while Chapter 6 focused on the Dutch group.

In **Chapter 5**, we calculated two sets of prosodic hyper-scores following Kalashnikova and Burnham (2018): general hyper-scores and unfamiliar hyper-scores. The general hyper-scores indicate to what extent IDS was generally exaggerated for each mother and the unfamiliar hyper-scores indicate how exaggerated IDS was specifically when mothers introduce unfamiliar words to children. Correspondingly, we asked two research questions:

> **RQ 5a**: Are there correlations between the generally exaggerated prosody of IDS (indicated by general hyper-scores) and children's vocabulary size?
> **RQ 5b**: Are there correlations between the prosody of IDS in word-learning contexts (indicated by unfamiliar hyper-scores) and children's vocabulary size?

For both Dutch and Mandarin Chinese, we investigated the concurrent correlations between the hyper-scores and children's vocabulary size. In addition, in order to examine whether IDS prosody at 18 months predicted children's vocabulary growth between 18 and 24 months, we examined the longitudinal correlations for Dutch in which we used vocabulary size at 18 months as one predictor.

The Dutch results revealed that there were no significant correlations between the general hyper-scores and children's vocabulary size, either concurrently or longitudinally. However, there were significant correlations between the Dutch unfamiliar hyper-scores and children's vocabulary size. Not only were the unfamiliar hyper-scores significantly correlated with children's vocabulary size concurrently at 18 and 24 months, they also significantly predicted children's vocabulary growth from 18 to 24 months of age. As for Mandarin Chinese, our results showed no significant correlations between the general hyper-scores and children's vocabulary size, nor were there any significant correlations between the unfamiliar hyper-scores and children's vocabulary size.

Chapter 6 reported a word-learning experiment (Experiment 3) which set out to address whether prototypical Dutch IDS facilitates 24-month-old children's word-to-object mapping compared to ADS (**RQ 6a**). This experiment was adapted from Ma et al. (2011) on American English, a study which is often cited as evidence for the facilitative effects of prototypical IDS on word learning. Our findings showed that Dutch 24-month-old children reliably learned novel word-to-object mappings in both the ADS and IDS conditions. There was only a small facilitative effect of IDS on one measure (mean single longest look), but overall our results did not strongly support the claim that prototypical IDS facilitates word learning. Based on the "hyper-score" measures in Chapter 5 and the word-learning results from Experiment 3, we further explored whether children's word learning performances in the ADS and IDS conditions were related to their input experience. Specifically, we asked whether the degree of prosodic exaggeration in IDS (indicated by the general hyper-scores) and whether the degree of prosodic exaggeration in IDS in word-learning contexts (indicated by the unfamiliar hyper-scores) significantly predicts children's performances in word-to-object mapping in ADS and IDS (**RQ 6b**). The results showed no significant correlations between the general hyper-scores or unfamiliar hyper-scores and children's word learning performances in IDS or ADS, at least for the small sample size.

The main goal of this dissertation was to investigate the role of prosodic input in word learning. To achieve this goal, we examined the nature of prosodic input in word-learning contexts, explored the correlations between individual prosodic input and children's language outcomes, and directly compared children's online word learning performances in the ADS and IDS conditions. The key findings of this dissertation, corresponding to the five research gaps, are summarized as follows. First, we asked what prosodic cues mothers used in word-learning contexts in Dutch and Mandarin Chinese IDS. We found that Dutch mothers had a lower mean pitch and a slower speaking rate specifically for unfamiliar words in IDS addressing 18- and 24-month-old children, while Mandarin Chinese mothers had a higher mean pitch specifically when they introduced unfamiliar words to 18-month-old children, and they had a larger pitch range in IDS specifically when introducing unfamiliar words to 24-month-old children. Second, we asked whether the prosody in word-learning contexts differs between Dutch and Mandarin Chinese. As illustrated above, we showed that the prosody of IDS and the prosody of IDS in word-learning contexts were manifested differently in these two languages (see Section 7.1.2 for more information). Third, we asked whether the prosody of IDS and the prosody of IDS in word learning contexts changed from 18 to 24 months of age. The results showed that for both languages, there was no evidence of age-related changes in speaking rate of IDS addressed to 18- and 24-month-old children, but the degree of pitch exaggeration does decrease in this period. Fourth, we asked whether prototypical IDS facilitates online word-to-object mapping for Dutch 24-month-old children. The results suggest that Dutch 24-month-old children can reliably learn novel word-to-object mappings from both ADS and IDS, however there is a slight advantage for the facilitative effects of IDS on online word learning. Finally, we examined whether there were correlations between the prosodic input in word-learning contexts and children's language outcomes. Our results revealed that the generally exaggerated prosody of IDS was not significantly correlated with children's vocabulary size, but the prosody of IDS specific to word-learning contexts was correlated with children's vocabulary size and vocabulary growth (at least for Dutch).

7.1.2 Cross-linguistic similarities and differences in IDS prosody

One aim of this dissertation was to better understand the language-universality and language-specificity of IDS prosody. So far, studies on IDS have usually examined a single language instead of taking a cross-linguistic perspective. As speech elicitation methods and speech content usually differ among the studies, it is difficult to make direct comparisons between their results. We therefore made two proposals concerning cross-linguistic investigations of IDS. First, since cross-linguistic investigations of IDS are valuable and necessary, we should use maximally similar speech elicitation methods in order to allow for a fair comparison between languages. Second, as IDS prosody (especially in word-learning contexts) is potentially language-specific, we should avoid making general statements about the effect of IDS on language acquisition based on findings from a single language. In accordance with these proposals, we conducted a cross-linguistic investigation of IDS in two typologically distinct languages (i.e., Dutch and Mandarin Chinese) using similar speech elicitation methods. This section summarizes the main findings from a cross-linguistic perspective and illustrates similarities and differences between Dutch and Mandarin Chinese IDS prosody. The language-universality and language-specificity of IDS will be discussed in more detail in Section 7.2.1.

Similarities between Dutch and Mandarin Chinese IDS prosody

First, both Dutch and Mandarin Chinese IDS were exaggerated in prosody when mothers addressed 18-month-old children. Specifically, IDS had a higher mean pitch than ADS at least until 18 months of age for both languages.

Second, IDS in both languages showed age-related changes in pitch from 18 months to 24 months, with a general trend of IDS prosody becoming more ADS-like as children grow older. Specifically, Mandarin Chinese IDS became ADS-like from 18 to 24 months of age in terms of pitch height and pitch range. The degree of pitch exaggeration in IDS was smaller in Dutch IDS addressing 24-month-old children compared to 18-month-old children.

Differences between Dutch and Mandarin Chinese IDS prosody

First, Dutch IDS addressed to 18- and 24-month-old children had a slower articulation rate than ADS, but Mandarin Chinese IDS did not show evidence of slowing down at either age.

Second, Mandarin Chinese IDS addressing 18-month-old children had a higher pitch and a larger pitch range than ADS, while IDS addressing 24-month-olds was similar to ADS with respect to both mean pitch and pitch range. Dutch IDS addressing 18- and 24-month-old children both had a higher pitch than ADS, while the pitch range did not differ between IDS and ADS.

Third, the prosodic input in word-learning contexts differed between Dutch and Mandarin Chinese. When Mandarin Chinese mothers addressed their 18-month-old children, utterance mean pitch increased specifically for unfamiliar words in IDS but not for familiar words. At 24 months, both word and utterance pitch range in IDS were exaggerated when mothers introduced unfamiliar words compared to familiar words. However, the Dutch results showed the opposite: Dutch mothers' word and utterance mean pitch raised specifically for familiar words instead of unfamiliar words for both age groups under investigation. Dutch mothers slowed down specifically for unfamiliar words in IDS.

Fourth, the prosody when introducing unfamiliar words was significantly correlated with children's vocabulary in Dutch, but similar evidence was not obtained from Mandarin Chinese.

7.2 General discussion and future directions

The main purpose of the current study was to investigate the role of prosodic input in children's word learning by examining Dutch and Mandarin Chinese IDS. This dissertation fills a number of research gaps in the literature and makes several contributions. First, it provides insights into the properties and possible functions of IDS in children's language acquisition. Specifically, it reveals the nature of prosodic input in word-learning contexts, extends previous literature on the effect of prototypical IDS on online word learning, and shows correlations between IDS prosody in word-learning contexts and children's vocabulary size for the first time. Second, this dissertation has implications for the language-universality and language-specificity of IDS. Finally, it focuses on IDS during children's vocabulary spurt period. By doing so, it contributes to our understanding of whether and how IDS might be adapted to different stages of language development. The remainder of this section will discuss these issues in more detail, provide directions for future research, and discuss methodological implications and challenges in IDS research.

7.2.1 Implications for the functions of IDS prosody

Researchers have proposed three functions of IDS: attracting infants' attention, conveying positive affect, and facilitating language acquisition (see reviews in Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015; Spinelli, Fasolo, & Mesman, 2017). Despite these proposals, a much-debated question is whether the prosody of IDS has particular linguistic functions. The attentional and affective functions of IDS have been shown to be related to the exaggerated prosody of IDS (Cooper & Aslin, 1994; Trainor, Austin, & Desjardins, 2000; but see Singh, Morgan, & Best, 2002). It is indeed possible that the exaggerated prosody of IDS may attract infants' attention to linguistic input and thus *indirectly* facilitates language acquisition. However, whether the prosody of IDS *directly* facilitates language acquisition beyond an attentional function is still a matter of debate.

Existing evidence for the linguistic functions of IDS is mostly related to the categorical learning of segmental (vowels and consonants) and suprasegmental features (lexical tones). Vowels in IDS are hyperarticulated compared to ADS in many languages, such as American English, Russian, Swedish, Taiwanese Mandarin, and French (Dodane & Al-Tamimi, 2007; Liu, Kuhl, & Tsao, 2003; Kuhl et al., 1997; though see Benders, 2013; Englund, 2017; and Miyazawa, Shinya, Martin, Kikuchi, & Mazuka, 2017 for opposing evidence). Lexical tones in IDS are found to be hyperarticulated in various tone languages, for instance, Cantonese, Northern Mandarin, and Taiwanese Mandarin (Liu, Tsao, & Kuhl, 2007; Tang et al., 2017, and Xu Rattanasone et al., 2013). In addition, Liu, Kuhl, and Tsao (2003) found that the degree of vowel hyperarticulation in Taiwanese Mandarin IDS addressing preverbal children (6–8 months and 10–12 months) is correlated with children's vowel discrimination skills. As such, researchers usually interpret the findings on vowel hyperarticulation and lexical tonal hyperarticulation as evidence that IDS enhances phonetic categorization by providing the optimal linguistic information for children (Cristia, 2013; Kuhl et al., 1997).

In addition to learning sound categories, children also learn words in the first two years of life. Despite the general agreement that vowel hyperarticulation and lexical tonal hyperarticulation in IDS may support phonetic categorization, the role of IDS in word learning is not yet clear. Indeed, vowel hyperarticulation may facilitate vocabulary development by facilitating phonetic categorization. For example, Kalashnikova and Burnham (2018) found that the degree of vowel hyperarticulation at 9 months was correlated with children's vocabulary size at a later age. In an online word recognition study, Song, Demuth, and Morgan (2010) found that vowel hyperarticulation in IDS could improve children's online word recognition compared to ADS. As Kalashnikova and Burnham (2018) pointed out, the mechanisms by which vowel hyperarticulation in IDS facilitates lexical development is not yet clear. However, as the major distinction between IDS and ADS is its prosody, a question that remains is: what is the role of IDS prosody in children's word learning?

In order to establish the role of IDS prosody in word learning, three issues should be considered. First, is the prosody of IDS organized in a way that may support word learning? Second, does the prosody of prototypical IDS facilitate children's online word learning compared to ADS? Third, are there any correlations between the prosody of IDS and children's word learning outcomes (vocabulary size and online word learning performances)?

To investigate the first issue, whether the prosody of IDS is varied such that it may support word learning, previous studies have examined the prosodic marking of contextually new information compared to contextually given information (Bortfeld & Morgan, 2010; Fernald & Mazzie, 1991; Fisher & Tokura, 1995). Also, mothers use prosody to distinguish adjectives (e.g., big vs. small) when addressing children (Herold, Nygaard, & Namy, 2012). Following the proposal by Fernald (2000), in addition to the prosodic modifications in IDS relevant to word recognition, it is possible that mothers would use prosody differently when they introduce familiar or unfamiliar words to children. However, no previous studies have specifically investigated whether mothers use prosody to highlight unfamiliar words compared to familiar words in IDS.

Second, recent evidence suggests that prototypical IDS facilitates children's online word processing (including word segmentation, word recognition, and word-to-object mapping) compared to ADS, but the evidence is largely based on American English IDS (Graf Estes & Hurley, 2013; Mani & Pätzold, 2016; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011; Singh, Nestor, Parikh, & Yull, 2009; Thiessen, Hill, & Saffran, 2005). In these studies, children's online word processing performances were tested when they heard audio stimuli that were produced in either prototypical IDS or ADS; the speech content of the two conditions was exactly the same and the only difference was prosody. Focusing on word-to-object mapping, in both Ma et al. (2011) and Graf Estes and Hurley (2013), children did not show a strong preference for IDS (indicated by longer looking time) in the training phase. Nevertheless, children performed better in the testing phase in the IDS condition compared to the ADS condition. As such, based on these two studies on American English, it seems that prototypical IDS prosody facilitates children's word-to-object mapping compared to ADS. However, the speech contexts in these studies were specifically word-learning contexts, in which mothers introduced unfamiliar words (usually pseudowords) to children. Consequently, it is not clear whether the facilitative effects of IDS found in these studies could in fact be attributed to the prosody of word-learning contexts or the generally exaggerated prosody of IDS. Moreover, as these studies have treated IDS as a package of features, including both prosodic as well as segmental features, it is unclear whether prosody can solely account for the facilitative effects of IDS on word learning. One study, however, specifically manipulated the effect of three factors—slow speaking rate, vowel hyperarticulation, and wide pitch range—of American English IDS on children's word recognition (Song et al., 2010). Their results showed that a slow speaking rate and vowel hyperarticulation, but not wide pitch range, significantly improved children's word recognition performance compared to ADS. Their findings suggest that the pitch properties of IDS do not necessarily serve to benefit language acquisition.

The results are mixed regarding the third issue, whether there are correlations between IDS prosody and children's language outcomes such as vocabulary size. A wider pitch range and a slower speaking rate in American English IDS have been shown to be correlated with a larger vocabulary size (Porritt, Zinser, Bachorowski, & Kaplan, 2014; Raneri, 2015). However, Kalashnikova and Burnham (2018) found that only vowel hyperarticulation, but not pitch level, was significantly correlated with children's vocabulary size. As such, it seems that in Australian English, only vowel hyperarticulation, but not generally exaggerated prosody, is related to the linguistic functions of IDS. However, as the authors pointed out, even though vowels in Australian English IDS are hyperarticulated, vowel *hypoarticulation* is found in many languages such as Dutch or Norwegian (Benders, 2013; Englund & Behne, 2005). As a result, whether their findings for Australian English would hold cross-linguistically is still unclear.

To summarize, so far it has been fundamentally unclear from previous studies whether the prosody of IDS has only attentional functions or has linguistic functions beyond attentional functions, especially in terms of word learning. The current dissertation may contribute to better understand this issue by taking two new perspectives:

First, our study is the first to investigate the prosody of IDS in word-learning contexts beyond its generally exaggerated prosody. We confirmed the familiarity of target words for each participant after testing. Our findings showed that mothers use different prosody when introducing unfamiliar and familiar words to their children in IDS compared to ADS. Specifically, Dutch mothers had a lower pitch and a slower articulation rate while Mandarin Chinese mothers tended to have a higher pitch and a larger pitch range when introducing unfamiliar words. These findings suggest that mothers do take their children's vocabulary knowledge into consideration and specifically distinguish familiar and unfamiliar words with prosody, which may potentially highlight unfamiliar words and thus facilitate their child's word learning.

Second, our study is the first to explore the correlations between prosodic input in word-learning contexts and children's vocabulary size. We found that the prosodic properties (i.e., pitch and speaking rate) specific to word-learning contexts *are* correlated with children's vocabulary size, whereas the prosodic properties of general IDS were not related to children's vocabulary size. Also, the F0 range of IDS at 18 months significantly predicted children's vocabulary growth between 18 and 24 months. These findings point to a possibility that was previously unconsidered: even though the generally exaggerated prosody of IDS may serve mainly attentional functions, IDS prosody specific to word-learning contexts might itself be fine-tuned to serve linguistic purposes. Thus, we argue that in order to better understand the role of prosodic input in children's word learning, it is necessary to investigate the nature of prosodic input specific to word-learning contexts in addition to examining the generally exaggerated prosody.

Even though we showed correlations between prosodic input and children's vocabulary size, we should interpret the significant correlations between prosodic input and children's vocabulary with caution. To begin with, it should be noted that these results are intrinsically correlational and cannot be used to address causality. We have not directly examined whether the prosodic cues that Dutch mothers use indeed facilitate children's word learning. As such, whether IDS prosody in word-learning contexts benefits children's novel word learning, and whether children rely on specific prosodic cues in word-learning contexts to learn novel words, requires further investigation. Furthermore, the correlations can be interpreted in two directions. In most studies that investigated the correlations between language input and children's language outcomes, significant correlations are often interpreted as evidence for the potential effect of language input on children's language outcomes. However, an alternative interpretation of the correlations would be that mothers *adapt* their IDS prosody in word-learning contexts precisely according to children's vocabulary knowledge. For instance, our Dutch results of the concurrent correlations at 18 months might indicate that mothers speak slower, have a lower word pitch, and reduce their F0 range when they introduce unfamiliar words to children with a relatively larger receptive vocabulary. Previous studies have shown that the prosody of IDS changes according to children's ages. For example, Kitamura et al. (2002) found that the mean pitch of Australian English IDS tended to increase from birth to 6 months, decreased at 9 months, and then increased at 12 months. In the same study, the authors found that the mean pitch of Thai IDS was highest at 9 months and had decreased at 12 months. Stern, Spieker, Barnett, and MacKain (1983) found that the pitch contour of IDS was most exaggerated at 4 months compared to newborns and 12- and 24-month-old children. In these studies, the prosody of IDS was often measured at the group level instead of the individual level. As there is considerable individual variation in children's vocabulary size (Bloom, 2002; Fenson et al., 2007), the results from these studies did not answer whether or how IDS might be adapted to individual children's vocabulary knowledge. Our findings on the concurrent correlations between IDS prosody and children's vocabulary size, however, suggest that mothers might adapt their IDS based on children's linguistic knowledge in a sophisticated way. However, our results on the longitudinal correlations revealed that F0 range at 18 months significantly predicted children's vocabulary growth between 18 and 24 months. These results could only be interpreted as evidence for the potential effect of IDS prosody (specifically F0 range) on children's language outcomes, as it is unlikely that a mother would adapt her speech according to her child's *future* vocabulary knowledge. Therefore, studies that examine correlations between IDS prosody and children's vocabulary growth will be able to make stronger claims on the role of prosodic input in children's word learning.

Third, the current study extends research on the facilitative effects of IDS on word-to-object mapping to Dutch and demonstrates inconsistent results with previous studies. Our findings suggest that Dutch 24-month-old children did not necessarily rely on IDS to learn novel word-to-object mapping. In fact, they could reliably learn novel word-to-object mapping in both the ADS and IDS conditions. Whether this inconsistency is due to differences in language, age groups, or experimental design (within-subject design vs. between-subject design) needs further investigation. As this line of research has not been carried out in languages other than American English or Dutch, whether prototypical IDS indeed facilitates children's online word learning should be tested in more languages. Additionally, determining to what age children are still dependent on IDS to learn novel words should also be tested in a longitudinal design.

Based on the discussion above, we propose that in order to understand the potential linguistic functions of IDS prosody, it is necessary to look beyond the generally exaggerated prosody of IDS and examine IDS prosody specific to word-learning contexts. Even though previous literature has suggested that the generally exaggerated prosody of IDS might be attentional or affective but not so much linguistic, our findings suggest that prosodic input specifically in word-learning contexts might be related to linguistic functions. Also, it is necessary to test whether prototypical IDS facilitates children's online word learning in a wider range of languages.

7.2.2 Language universality and language specificity in IDS

Ever since Ferguson's (1978) pioneering study on IDS entitled "Talking to children: a search for universals," IDS studies have been devoted to providing evidence for the language-universality of IDS. To date, researchers have demonstrated that caregivers around the world modify their prosody when talking to children. The exaggerated speaking style of IDS has been found in a wide range of languages including American English, British English, Australian English, Cantonese, French, German, Japanese, Korean, Mandarin Chinese, Sri Lankan Tamil, Thai, Tagalog, to name a few (Fernald et al., 1989; Grieser & Kuhl, 1988; Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002; Narayan & McDermott, 2016; Xu, 2008). So far, only one language has shown counterevidence, that is, Quiché Mayan (Ingram, 1995; Bernstein Ratner & Pye, 1984). As a result, researchers have reached an agreement that the exaggerated prosody of IDS is universal across languages and cultures (Cristia 2013; Soderstrom, 2017) and assume "only slight differences across languages and cultures" in IDS (Spinelli, Fasolo, & Mesman, 2017, p. 2). The "slight differences" particularly refer to the *degree* of prosodic exaggeration. For example, the difference in mean pitch between ADS and IDS is larger in American English than in British English, French, Italian, German, or Japanese (Fernald et al., 1989). Also, Australian English IDS is more exaggerated in mean pitch and pitch range than Thai IDS compared to ADS (Kitamura et al., 2002). Taken together, previous studies suggest that prosodic exaggeration in IDS is universal, with only the degree of prosodic exaggeration varying across languages.

Even though we found evidence for overall prosodic exaggeration in both Dutch and Mandarin Chinese, consistent with the language-universality of IDS prosody, our findings also provide evidence for the language-specificity of IDS prosody. Cross-linguistic differences in IDS prosody are *not* limited to the degree of prosodic exaggeration. In particular, the specific prosodic cues that were exaggerated differ between Dutch and Mandarin Chinese IDS, the age-related changes of IDS prosody are manifested differently in Dutch and Mandarin Chinese, and importantly, the prosody in word-learning contexts differs between these two languages (see Section 7.1.2).

The prosody of IDS has been found to be generally exaggerated in a wide range of languages with diverse types of word prosody, for example, tonal languages (e.g., Mandarin Chinese, Thai, and Cantonese) and stress languages (e.g., English and Dutch), however, IDS prosody in word-learning contexts may differ. Our findings suggest that Dutch and Mandarin Chinese mothers may use different prosody when introducing unfamiliar words to children, which may be related to the native language's word prosody (lexical stress vs. lexical tone) and rhythmic property (stress-timed vs. syllable-timed). Consequently, as the prosody in word-learning contexts varies across languages with different prosodic characteristics, the specific cues that may account for the potential linguistic functions may also differ among languages. However, whether the findings in our studies can be attributed to prosodic typology, and how prosodic typology may play a role in IDS prosody, needs further investigation. As such, we suggest that future research compare IDS in both prosodically similar and prosodically different languages to determine the effect of prosodic typology on the prosody of IDS. For example, it would be interesting to examine whether stress languages such as English or German also mainly use a slower speaking rate instead of a heightened pitch when introducing unfamiliar words in IDS, and whether tonal languages such as Thai or Cantonese also use pitch to highlight unfamiliar words in IDS. One language that would make for a particularly interesting study is Ambonese Malay, which has no lexical stress, no lexical tones, and no prosodic focus (Maskikit-Essed & Gussenhoven, 2016). As pitch seems not to be used for word prosody or pragmatic purposes in this language, one direction for future research is to examine which cues Ambonese-Malay-speaking mothers may use to highlight unfamiliar words in IDS, for example, speaking rate, pausing, and/or word order. In sum, the cross-linguistic differences of IDS prosody suggest that prosodic typology should be taken into consideration when investigating IDS prosody.

In addition to the cross-linguistic differences in IDS prosody, the effect of IDS on word learning may differ among languages as well. Previous studies suggest that prototypical IDS facilitates American-English-learning children's online word-to-object mapping at least until 27 months (Graf Estes & Hurley, 2013; Ma et al., 2011). However, we conducted an adapted experiment from Ma et al. (2011) for Dutch and found that Dutch 24-month-old children could reliably learn novel words in both the ADS and IDS conditions. The differences between our results and Ma et al. (2011) can possibly be attributed to the different ages under investigation or the different experimental designs (between-subject design vs. within-subject design). However, it is equally possible that the facilitative effects of prototypical IDS that hold for American English may not necessarily extend to other languages such as Dutch.

Taken together, the language-specificity of IDS should be further investigated for a larger set of languages, taking into account typological differences in prosody. Also, whether prototypical IDS facilitates word learning across languages will require further investigation on more languages. Thus, future research may investigate two questions. First, does prototypical IDS indeed facilitate children's online word learning compared to ADS across languages? Second, is prototypical IDS necessary for children to learn novel words across languages?

7.2.3 IDS during the "vocabulary spurt" period

Most studies comparing the prosody of IDS to ADS have focused on the first year of life, and the implications of these studies have been mainly related to attracting infants' attention or pre-linguistic skills such as eliciting response and imitation of parents' voices (Fernald, 1989; Spinelli et al., 2017). Several studies also investigated the age-related changes of IDS prosody in the first year of life, and there is some evidence to suggest that the prosody of IDS in the first year of life is most exaggerated compared to later ages (Kitamura et al., 2002; Stern, Spieker, Barnett, & MacKain; 1983; Liu, Tsao, & Kuhl, 2009). If the prosody of IDS only serves to benefit infants before one year of age, we would predict that IDS becomes similar to ADS when children are older than 12 months. In the literature, IDS during the second year of life is much less studied compared to IDS in the first year. In a meta-analysis on the relationship between IDS prosody and children's language outcomes, only one out of fifteen studies investigated IDS addressed to children at 20 months of age, while the rest had all focused on IDS addressed to infants before 12 months (Spinelli et al., 2017). Thus, it is unclear from the literature what the role of IDS prosody is in the second year, especially during the "vocabulary spurt" period (around 16 to 18 months to 24 months) when children's vocabulary size significantly increases and their "fast mapping" ability gradually improves (Bion et al., 2013; Goldfield & Reznick, 1990). Would IDS still be exaggerated during this period? Also, if IDS adapts to different stages of language development, would IDS in this period show characteristics that may benefit vocabulary learning?

As we were specifically interested in the role of IDS in word learning, we selected 18 months and 24 months as our target age groups. The findings on this age group provide insight into the prosodic input in children's vocabulary spurt period. We found that the prosodic exaggeration of IDS continues well into the second year of life for both Dutch and Mandarin Chinese. Also, IDS in both languages showed age-related changes in pitch from 18 months to 24 months, and a general trend is that IDS prosody gradually becomes ADS-like as children grow older. These findings suggest that the prosody of IDS becomes similar to ADS as children's vocabulary size increases. Crucially, we found that a larger F0 range of individual Dutch IDS at 18 months significantly predicted children's vocabulary growth between 18 and 24 months, suggesting that the quality of IDS prosody plays a role in children's vocabulary spurt.

7.2.4 Methodological issues

We propose that it is important to use maximally similar speech elicitation methods in cross-linguistic investigations on IDS and use similar speech contexts in ADS and IDS conditions. In previous studies, the speech elicitation methods vary. In particular, three speech elicitation methods are commonly used for the IDS condition: (1) natural mother-child interaction at home (e.g., Fernald et al., 1989; Kitamura et al., 2002; Narayan & McDermott, 2016); (2) semi-spontaneous speech elicitation with target objects/books (e.g., Burnham et al., 2015; Herold, Nygaard, & Namy, 2012; Liu, Tsao, & Kuhl, 2007; Kalashnikova & Burnham, 2018); and (3) scripted read speech (e.g., Ko & Soderstrom, 2013; Shute & Wheldall, 1999). In the ADS condition, the participant mothers usually have a free conversation with an experimenter. The first method ensures the naturalness of the speech data, however, it should be noted that the content of IDS differs depending on the various daily activities and different contexts involved (Tamis-LeMonda, Custode, Kuchirko, Escobar, & Lo, 2018). Consequently, the prosody of IDS may differ to a large degree among the different activities or contexts. As a result, the speech contexts are not comparable between the ADS and IDS conditions if natural home recordings are used. In this dissertation, we used a semi-spontaneous storybook-telling task for both the ADS and IDS conditions in order to set up word-learning contexts, and we had similar speech elicitation methods and materials for Dutch and Mandarin Chinese. This methodological choice allows us to establish a similar word-learning context in each language and to have more comparable speech contexts between conditions. However, it is important to note that our conclusions from the cross-linguistic perspective are drawn from *qualitative* comparisons of the results from each language, rather than *quantitative* analyses that included the two languages in the same models.

Despite the advantages of using similar speech elicitation methods across speech registers (IDS condition vs. ADS condition) and across languages, our experimental design has its limitations. First, it is not clear whether the prosody during shared book-reading, even if the speech is not read speech, is representative of word-learning contexts in natural mother-child interactions at home. Second, we had mothers read a picture book to another adult in the ADS condition, which is not typical of daily life and may have led to unnatural speech. Third, another limitation of the current study is that we had different experimental designs between Dutch and Mandarin Chinese IDS. Specifically, we used a longitudinal design for Dutch but a cross-sectional design for Mandarin Chinese. This was largely due to practical issues in data collection. As we did not find any differences in ADS between the two age groups in Mandarin Chinese in Chapters 2, 3, and 4, the difference in design is unlikely to affect our interpretations of results. Fourth, we only evaluated Dutch children's online word learning performances. A further study may investigate whether the prosody of Mandarin Chinese changed longitudinally and whether prototypical Mandarin Chinese IDS facilitates children's online word learning. Finally, the correlation results reported in chapters 5 and 6 were limited in the sense that sample sizes were rather small. Future studies may explore the correlations between IDS prosody and children's language outcomes by including larger sample sizes than we were able to.

7.2.5 A future direction: corpora of Dutch and Mandarin Chinese IDS

A future direction of the current dissertation is to further annotate the audio and video data collected for this dissertation and to build open access corpora of Dutch and Mandarin Chinese IDS in order to make the data available for other researchers.

The current dissertation has reported results on word and utterance prosody in Dutch and Mandarin Chinese as well as lexical tonal prosody in Mandarin Chinese. Further annotation of the data will certainly allow researchers to examine different aspects of the speech data from acoustical, phonological, syntactic, and pragmatic points of view. For example, possible topics include other segmental or prosodic cues in the speech data, such as disfluency and pausing in IDS, word stress in Dutch IDS, vowel quality in IDS, and prosodic phrasing as a cue to syntactic structure, to name a few.

An open corpus would also allow investigation of multimodal aspects of language input, including speech and gestures, and their temporal alignment. Compared to the large amount of literature on IDS, relatively few studies have examined its accompanying maternal gestures (e.g., pointing and touching). There is a growing recognition of the role of maternal gestures in children's language acquisition (e.g., Gogate, Bahrick, & Watson, 2000; Olson & Masur, 2015; O'Neill, Bard, Linnell, & Fluck, 2005; Vallotton, Decker, Kwon, Wang, & Chang, 2017). Further annotation of the multimodal aspects of the video data collected for this dissertation would add to this research, specifically addressing word-learning contexts in Dutch and Mandarin mother-child interactions.

7.3 Conclusion

The current dissertation set out to investigate the role of prosodic input in children's word learning by examining infant-directed speech in Dutch and Mandarin Chinese—two typologically distinct languages. This research has three key findings: first, prosodic input in word-learning contexts differs between Dutch and Mandarin Chinese. Thus, even though IDS has generally exaggerated prosody across languages, the prosody of IDS in word-learning contexts is language-specific. Second, IDS prosody in word-learning contexts in Dutch, rather than its generally exaggerated prosody, is correlated with children's vocabulary size; Third, 24-month-old Dutch children do not rely on prototypical IDS to learn novel words, but IDS might have a small facilitative effect on online word learning compared to ADS. These findings together suggest that mothers are sensitive, consciously or unconsciously, to their children's vocabulary knowledge and adapt their speech prosody accordingly when they introduce unfamiliar words to children. Thus, understanding the nature of prosodic input in word-learning contexts is a crucial step towards understanding the role of IDS in word learning. In conclusion, the findings reported here shed new light on the role of prosodic input in language acquisition: the generally exaggerated prosody of IDS may be mainly aimed at drawing infants' attention, but the prosodic input in word-learning contexts is fine-tuned for linguistic purposes and may play a significant role in children's early word learning.

References

Abercrombie, D. (1967). *Elements of general phonetics*. Chicago, IL: Aldine Publishing Company.

Adriaans, F., & Swingley, D. (2017). Prosodic exaggeration within infant-directed speech: Consequences for vowel learnability. *The Journal of the Acoustical Society of America*, *141*(5), 3070–3078. https://doi.org/10.1121/1.4982246

Bailey, T., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, *17*(2), 1267–1284.

Baker, R. E., & Bradlow, A. R. (2009). Variability in word duration as a function of probability, speech style, and prosody. *Language and Speech*, *52*(4), 391–413. https://doi.org/10.1177/0023830909336575

Bao, Z. (1996). The syllable in Chinese. *Journal of Chinese Linguistics, 24*(2), 312–354.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*. 68, 255–278. doi: 10.1016/j.jml.2012.11.001

Barry, J. G., & Blamey, P. J. (2004). The acoustic analysis of tone differentiation as a means for assessing tone production in speakers of Cantonese. *The Journal of the Acoustical Society of America*, *116*(3), 1739–1748.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. Available at: http://arxiv.org/abs/1506.04967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01.

Beckman, M. E. (1992). Evidence for speech rhythm across languages. In: Y. Tohkura, E. Vatikiotis-Bateson & Y. Sagisaka (Ed.) *Speech perception, production and linguistic structure* (pp. 457–463). Amsterdam, The Netherlands: IOS Press.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*(1), 92–111.

Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, *36*(4), 847–862. https://doi.org/10.1016/j.infbeh.2013.09.001

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of*

Sciences, *109*(9), 3253–3258.

Bernstein Ratner, N., & Pye, C. (1984). Higher pitch in BT is not universal: Acousticevidence from Quiché Mayan. *Journal of Child Language*, *11*(3)., 515–522. https://doi.org/10.1017/S0305000900005924

Bernstein-Ratner, N. (1985). Dissociations between vowel durations and formant frequency characteristics. *Journal of Speech, Language, and Hearing Research*, *28*(2), 255–264.

Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*(1), 39–53. https://doi.org/10.1016/j.cognition.2012.08.008

Blicher, D. L., Diehl, R. L., & Cohen, L. B. (1990). Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics, 18*(1), 37–49.

Bloom, P. (2001). Précis of *How children learn the meanings of words*. *Behavioral and Brain Sciences*, *24*(06), 1095–1103.

Bloom, P. (2002). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Boersma, P., & Weenink, D. J. M. (2017). Praat: Doing phonetics by computer [Computer program]. Available at: http://www.praat.org/.

Bortfeld, H., & Morgan, J. L. (2010). Is early word-form processing stress-full? How natural variability supports recognition. *Cognitive Psychology*, *60*(4), 241–266.

Brooks, P. J., & Kempe, V. (2014). *Encyclopedia of language development*. Thousand Oaks, CA: Sage Publications.

Bryant, G. A., & Barrett, H. C. (2007). Recognizing intentions in infant-directed speech evidence for universals. *Psychological Science*, *18*(8), 746–751.

Burnham, E. B., Wieland, E. A., Kondaurova, M. V., McAuley, J. D., Bergeson, T. R., & Dilley, L. C. (2015). Phonetic modification of vowel space in storybook speech to infants up to 2 Years of Age. *Journal of Speech Language and Hearing Research*, *58*(2), 241–253. https://doi.org/10.1044/2015_JSLHR-S-13-0205

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS one*, *5*(6), e10729.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: C. Li (Ed.) *Subject and topic* (pp. 25–55). New York, NY: Academic Press.

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., Ford, C., Volcic, R., & De Rosario, M. H. (2018). Package ‘pwr’. Retrieved from https://cran.r-project.org/web/packages/pwr/pwr.pdf.

Chao, Y. R. (1948) *Mandarin primer*. Cambridge, MA: Harvard University Press.

Chen, A. (2009). The phonetics of sentence-initial topic and focus in adult and child Dutch. In M. C. Vigario, S. Frota & M. João Freitas (Ed.), *Phonetics and phonology: Interactions and interrelations* (pp. 91–106). Amsterdam, The Netherlands: John Benjamins.

Chen, A., Stevens, C. J., & Kager, R. (2017). Pitch perception in the first year of life, a comparison of lexical tones and musical pitch. *Frontiers in Psychology*, *8*, 297.

Chen, S.-W., Wang, B., & Xu, Y. (2009). Closely related languages, different ways of realizing focus. In *Proceedings of the 10th Annual Conference of the International Speech Communication Association* (Interspeech 2009) (pp. 1007–1010), Brighton, UK.

Chen, Y., & Braun, B. (2006). Prosodic realization of information structure categories in Standard Chinese. In R. Hoffmann, H. Mixdorff (Ed.), *Proceedings of the International Conference on Speech Prosody.* Dresden, Germany: TUD.

Chen, Y., & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, *36*(4), 724–746.

Church, R., Bernhardt, B., Shi, R., & Pichora-Fuller, K. (2005). Infant-directed speech: Final syllable lengthening and rate of speech. *The Journal of the Acoustical Society of America*, *117*(4), 2429–2430. https://doi.org/10.1121/1.4786663

Cohen Priva, U. (2017). Not so fast: Fast speech correlates with lower lexical and structural information. *Cognition,* 160. 27–34.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed Speech in the first month after birth. *Child Development*, *61*(5), 1584–1595. https://doi.org/10.1111/j.1467-8624.1990.tb02885.x

Cooper, R. P., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development*, *65*(6), 1663–1677. https://doi.org/10.1111/j.1467-8624.1994.tb00841.x

Cristia, A. (2013). Input to Language: The phonetics and perception of infant-directed Speech. *Language and Linguistics Compass*, *7*(3), 157–170. https://doi.org/10.1111/lnc3.12015

Cristia, A., & Seidl, A. (2014). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language*, *41*(4), 913–934.

De Mooij, A.J. (2018). The UiL OTS Video Coding System (Version 2) [Computer software]. Utrecht Institute of Linguistics OTS, Utrecht University. Available at: https://github.com/UiL-OTS-labs-backoffice/UiL-OTS-Video-Coding-System

De Ruiter, L. E. (2015). Information status marking in spontaneous vs. read speech in story-telling tasks–Evidence from intonation analysis using GToBI. *Journal of Phonetics*, *48*, 29–44.

Delle Luche, C., Durrant, S., Poltrock, S., & Floccia, C. (2015). A methodological investigation of the Intermodal Preferential Looking paradigm: Methods of analyses, picture selection and data rejection criteria. *Infant Behavior and Development*, *40*, 151–172. https://doi.org/10.1016/j.infbeh.2015.05.005

Dodane, C., & Al-Tamimi, J. (2007). An acoustic comparison of vowel systems in adult-directed speech and child-directed speech: Evidence from French, English and Japanese. Paper presented at the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.

Duanmu, S. (2007). *The phonology of Standard Chinese* (2nd ed.). New York, NY: Oxford University Press.

Dunst, C., Gorman, E., & Hamby, D. (2012). Preference for infant-directed speech in preverbal young children. *Center for Early Literacy Learning*, *5*(1), 1–13.

Englund, K. T. (2017). Hypoarticulation in infant-directed speech. *Applied Psycholinguistics*, 1–21. https://doi.org/10.1017/S0142716417000480

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., & Reznick, J. S. (2007). *MacArthur-Bates communicative development inventories: User's guide and technical manual*. Baltimore, MD: Brookes.

Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist*, *66*, 103–114. https://doi.org/10.1525/aa.1964.66.suppl_3.02a00060

Ferguson, C. A. (1978). Talking to children: a search for universals. In J. Greenberg, C. A. Ferguson & E. A. Moravcsik (Ed.), *Universals of human language. Volume I: method and theory* (pp. 203–224). Palo Alto, CA: Stanford University Press.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, *8*(2), 181–195.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, *60*(6), 1497–1510. https://doi.org/10.2307/1130938

Fernald, A. (2000). Speech to infants as hyperspeech: Knowledge-driven processes in early word recognition. *Phonetica*, *57*, 242–254.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults.

Developmental Psychology, *27*(2), 209–221. https://doi.org/10.1037/0012-1649.27.2.209

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, *64*(3), 637–656.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104–113. https://doi.org/10.1037/0012-1649.20.1.104

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477–501. https://doi.org/10.1017/S0305000900010679

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E. M. Fernández & H. Clahsen (Ed.), *Developmental psycholinguistics: Online methods in children's language processing* (pp. 97–135). Amsterdam, The Netherlands: John Benjamins.

Fisher, C., & Tokura, H. (1995). The given-new contract in speech to infants. *Journal of Memory and Language*, *34*(3), 287–310. https://doi.org/10.1006/jmla.1995.1013

Floccia, C., Keren-Portnoy, T., DePaolis, R., Duffy, H., Delle Luche, C., Durrant, S., … Vihman, M. (2016). British English infants segment words only with exaggerated infant-directed speech stimuli. *Cognition*, *148*, 1–9. https://doi.org/10.1016/j.cognition.2015.12.004

Frank, M. C., Bergelson, E., Bergmann, C., Cristia, A., Floccia, C., Gervain, J., ... & Lew-Williams, C. (2017). A collaborative approach to infant research: Promoting reproducibility, best practices, and theory-building. *Infancy*, *22*(4), 421–435.

Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, *11*(1), 43–79.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 878–894.

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*(1), 171–183.

Golinkoff, R. M., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language*, *22*(3), 703–726. https://doi.org/10.1017/S0305000900010011

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to me the social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science*, *24*(5), 339–344. https://doi.org/10.1177/0963721415595345

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Ed.), *Papers in laboratory phonology* (pp. 515–546). Berlin: Mouton de Gruyter.

Graf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy*, *18*(5), 797–824. https://doi.org/10.1111/infa.12006

Grassmann, S., & Tomasello, M. (2007). Two-year-olds use primary sentence accent to learn new words. *Journal of Child Language*, *34*(3), 677–687.

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, *24*(1), 14–20.

Gundel, J. K. (1999). Topic, focus, and the grammar-pragmatics interface. In: Alexander, N.J., Minnick, M., (Eds.), *Proceedings of the 23rd Annual Penn Linguistics Colloquium, vol. 6.1. Penn Working Papers in Linguistics,* (pp. 185–200). Available at: https://repository.upenn.edu/pwpl/vol6/iss1/14

Halliday, M. A. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, *3*(2), 199–244.

Han, M., de Jong, N. H., & Kager, R. (2018a). Lexical tones in Mandarin Chinese infant-directed speech: Age-related changes in the second year of life. *Frontiers in Psychology*, *9*, 434.

Han, M., de Jong, N. H., & Kager, R. (2018b). Infant-directed speech is not always slower: Cross-linguistic evidence from Dutch and Mandarin Chinese. In A. Bertolini and M. Kaplan (Ed.) *Proceedings of the 42nd annual Boston University Conference on Language Development* (pp. 331–344). Cascadilla Press.

Hanssen, J., Peters, J., & Gussenhoven, C. (2008). Prosodic effects of focus in Dutch declaratives. In P. Barbosa, S. Madureira, & C. Reis (Ed.): *Proceedings of the 4th Conference on Speech Prosody* (pp. 423–426). Campinas, Brazil: Editoria RG/CNPq.

Harrell, F. (2018). Hmisc: Harrell Miscellaneous. R package version 4.1-1. Available at: https://CRAN.R-project.org/package=Hmisc.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H Brookes Publishing.

Hartman, K. M., Ratner, N. B., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language*, *44*(5), 1140–1162.

Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech Language and Hearing Research*, *44*(6), 1189–1200. https://doi.org/10.1044/1092-4388(2001/092)

Hermes, D. J., & Van Gestel, J. C. (1991). The frequency scale of speech intonation. *The Journal of the Acoustical Society of America*, *90*(1), 97–102.

Herold, D. S., Nygaard, L. C., & Namy, L. L. (2012). Say it like you mean it: mothers' use of prosody to convey word meaning. *Language and Speech*, *55*(3), 423–436. https://doi.org/10.1177/0023830911422212

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (Ed.), *Methods for assessing children's syntax* (pp. 105–124). Cambridge, MA: MIT Press.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*(2), 418–433. https://doi.org/10.1111/1467-8624.00415

Hoole, P., & Hu, F. (2004, March). Tone-vowel interaction in Standard Chinese. Paper presented at the International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages, Beijing, China.

Howie, J. M. (1974). On the domain of tone in Mandarin. *Phonetica*, *30*(3), 129–148.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*(6), F31–F39. https://doi.org/10.1111/j.1467-7687.2008.00768.x

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Ingram, D. (1995). The cultural basis of prosodic modifications to infants and children: a response to Fernald's universalist theory. *Journal of Child Language*, *22*(1), 223–233. https://doi.org/10.1017/S0305000900009715

Johnson, E. K., Lahey, M., Ernestus, M., & Cutler, A. (2013). A multimodal corpus of speech to infant and adult listeners. *The Journal of the Acoustical Society of America*, *134*(6), EL534–EL540. https://doi.org/10.1121/1.4828977

Jusczyk, P. W., & Aslin, R. N. (1995). Infants′ detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*(1), 1–23. https://doi.org/10.1006/cogp.1995.1010

Kalashnikova, M., & Burnham, D. (2018). Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *Journal of Child Language*, 1–19. https://doi.org/10.1017/S0305000917000629

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*(3), 643–650.

Kitamura, C., & Burnham, D. (1998). The infant's response to maternal vocal affect. *Advances in Infancy Research*, *12*, 221–236.

Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy*, *4*(1), 85–110.

Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2002). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, *24*(4), 372–392. https://doi.org/10.1016/S0163-6383(02)00086-3

Knoll, M. A., & Costall, A. (2015). Characterising F(0) contour shape in infant- and foreigner-directed speech. *Speech Communication*, *66*, 231–243. https://doi.org/10.1016/j.specom.2014.10.007

Ko, E.-S., & Soderstrom, M. (2013). Additive effects of lengthening on the utterance-final word in child-directed speech. *Journal of Speech, Language, and Hearing Research*, *56*(1), 364–371. https://doi.org/10.1044/1092-4388(2012/11-0341)

Krifka, M, Musan, R (2012) Information structure: Overview and linguistic issues. In: M. Krifka, & R. Musan. (Ed.) *The expression of information structure*: *Volume: The Expression of Cognitive Categories (ECC)* (pp. 1–44). Berlin: Mouton de Gruyter.

Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, Stolyarova, E.I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*, 684–686.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). 1–26. doi: 10.18637/jss.v082.i13

Ladefoged, P. (1975). *A course in phonetics*. New York: Harcourt Brace Jovanovich.

Lennes, M. (2017). SpeCT-Speech Corpus Toolkit for Praat (v1.0.0). First release on GitHub. Retrieved from https://github.com/lennes/spect/releases

Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2017). The temporal modulation structure of infant-directed speech. *Open Mind*, *1*(2), 78–90.

Li, D. C., & Lee, S. (2006). Bilingualism in East Asia. In Bhatia, T. K., and Ritchie, W. C., (Ed.) *The handbook of bilingualism* (pp. 742–779), Malden, MA: Blackwell Publishing Ltd. doi: 10.1002/9780470756997.ch28

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and speech*, *6*(3), 172–187.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Ed.) *Speech production and speech modelling*. (pp. 403–439). Dordrecht: Kluwer.

Liu, H.-M., Kuhl, P. K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), F1–F10. https://doi.org/10.1111/1467-7687.00275

Liu, H.-M., Tsao, F.-M., & Kuhl, P. K. (2007). Acoustic analysis of lexical tone in Mandarin infant-directed speech. *Developmental Psychology*, *43*(4), 912–917. https://doi.org/10.1037/0012-1649.43.4.912

Liu, H.-M., Tsao, F.-M., & Kuhl, P. K. (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: a longitudinal study. *Journal of Child Language*, *36*(4), 909–922. https://doi.org/10.1017/S030500090800929X

Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, *133*(2), 385–394.

Liu, L., & Kager, R. (2017). Perception of tones by bilingual infants learning non-tone languages. *Bilingualism: Language and Cognition*, *20*(3), 561–575.

Liu, L., & Kager, R. (2018). Monolingual and bilingual infants' ability to use non-native tone for word learning deteriorates by the second year after birth. *Frontiers in Psychology*, *9*, 117. https://doi.org/10.3389/fpsyg.2018.00117

Ma, W., Golinkoff, R. M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development*, *7*(3), 185–201.

Mahr, T., & Edwards, J. (2018). Using language input and lexical processing to predict vocabulary size. *Developmental Science*, *21*(6), e12685. https://doi.org/10.1111/desc.12685

Mani, N., & Pätzold, W. (2016). Sixteen-month-old infants' segment words from infant- and adult-directed speech. *Language Learning and Development*, *12*(4), 499–508. https://doi.org/10.1080/15475441.2016.1171717

Martin, A., Igarashi, Y., Jincho, N., & Mazuka, R. (2016). Utterances in infant-directed speech are shorter, not slower. *Cognition*, *156*, 52–59. https://doi.org/10.1016/j.cognition.2016.07.015

Masataka, N. (1992). Pitch characteristics of Japanese maternal speech to infants. *Journal of Child Language*, *19*(2), 213–223.

Maskikit-Essed, R., & Gussenhoven, C. (2016). No stress, no pitch accent, no prosodic focus: the case of Ambonese Malay. *Phonology*, *33*(2), 353–389. https://doi.org/10.1017/S0952675716000154

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, *10*(3), 241–265.

Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, *106*(3), 1367–1381.

McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence?. *Cognition*, *129*(2), 362–378.

Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H., & Mazuka, R. (2017). Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition*, *166*, 84–93. https://doi.org/10.1016/j.cognition.2017.05.003

Mok, P., & Dellwo , V. (2008). Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. In P. Barbosa, S. Madureira, & C. Reis (Ed.), *Proceedings of the 4th Conference on Speech Prosody* (pp. 423–426). Campinas, Brazil: Editoria RG/CNPq

Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustical Society of America*, *139*(3), 1272–1281.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: two modes of word acquisition? *Developmental Science*, *6*(2), 136–142. https://doi.org/10.1111/1467-7687.00263

Newman, R. S., Rowe, M. L., & Bernstein Ratner, N. (2015). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, *43*(5), 1158–1173. https://doi.org/10.1017/S0305000915000446

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38.

Olson, J., & Masur, E. F. (2015). Mothers' labeling responses to infants' gestures predict vocabulary outcomes. *Journal of Child Language*, *42*(6), 1289–1311.

O'Neill, M., Bard, K. A., Linnell, M., & Fluck, M. (2005). Maternal gestures with 20-month-old infants in two contexts. *Developmental Science*, *8*(4), 352–359.

Papoušek, M., & Hwang, S. F. C. (1991). Tone and intonation in Mandarin babytalk to presyllabic infants: Comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics*, *12*(4), 481–504.

Porritt, L. L., Zinser, M. C., Bachorowski, J.-A., & Kaplan, P. S. (2014). Depression diagnoses and fundamental frequency-based acoustic cues in maternal infant-directed speech. *Language Learning and Development*, *10*(1), 51–67. https://doi.org/10.1080/15475441.2013.802962

Pye, C. (1986). Quiché Mayan speech to children. *Journal of Child Language*, *13*(1), 85–100. https://doi.org/10.1017/S0305000900000313

Quené, H. (2007). On the just noticeable difference for tempo in speech. *Journal of Phonetics*, *35*(3), 353–362.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, *17*(6), 880–891.

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: perceptual studies. Paper presented at the 15th International Congress of Phonetic Sciences, Barcelona, Spain.

Raneri, D. P. (2015). *Infant-directed speech: Maternal pitch variability, rate of speech, and child language outcomes*. (Master's thesis, University of Maryland, College Park). Retrieved from https://drum.lib.umd.edu/handle/1903/16736

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child* *development*, *83*(5), 1762–1774.

Saxton, M. (2008). What's in a name? Coming to terms with the child's linguistic environment. *Journal of Child Language*, *35*(3), 677–686.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by 15-month-olds under tightly controlled conditions. *Child Development*, *69*(2), 309–320.

Segal, J., & Newman, R. S. (2015). Infant preferences for structural and prosodic properties of infant-directed speech in the second year of life. *Infancy*, *20*(3), 339–351.

Seifart, F., Strunk, J., Danielsen, S., Hartmann, I., Pakendorf, B., Wichmann, S., ... & Bickel, B. (2018). Nouns slow down speech across structurally and culturally diverse languages. *Proceedings of the National Academy of Sciences*, *115*(22), 5720–5725.

Shute, B., & Wheldall, K. (1999). Fundamental frequency and temporal modifications in the speech of British fathers to their children. *Educational Psychology*, *19*(2), 221–233. https://doi.org/10.1080/0144341990190208

Shute, B., & Wheldall, K. (2001). How do grandmothers speak to their grandchildren? Fundamental frequency and temporal modifications in the speech of British grandmothers to their grandchildren. *Educational Psychology*, *21*(4), 493–503. https://doi.org/10.1080/01443410120090858

Singh, L., & Fu, C.S.L. (2016). A new view of language development: The acquisition of lexical tone. *Child Development, 87*(3)*,* 834–854.

Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, *3*(3), 365–394. https://doi.org/10.1207/S15327078IN0303_5

Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy*, *14*(6), 654–666.

Sjons, J., Hörberg, T., Östling, R., & Bjerva, J. (2017). Articulation rate in Swedish child-directed speech increases as a function of the age of the child even when surprisal is controlled for. In: M. Włodarczak (Ed.) *Proceedings of the 18th Annual Conference of the International Speech Communication Association* (pp. 1794–1798).

Radek, S., & Machač, P. (2011). *Principles of phonetic segmentation*. Prague, Czech: Epocha Publishing House.

Sluijter, A. M., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical society of America*, *100*(4), 2471–2485.

Snow, C. E. (1977). Mothers' speech research: From input to interaction. In C. E.

Snow & C. A. Ferguson (Ed.), *Talking to children* (pp. 31–50). New York: Cambridge University Press.

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501–532. https://doi.org/10.1016/j.dr.2007.06.002

Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J. L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, *35*(4), 869–902.

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America*, *128*(1), 389–400. https://doi.org/10.1121/1.3419786

Song, J. Y., Demuth, K., & Morgan, J. (2018). Input and processing factors affecting infants' vocabulary size at 19 and 25 months. *Frontiers in Psychology*, *9*, 2398. https://doi.org/10.3389/fpsyg.2018.02398

Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review*, 44, 1–18. https://doi.org/10.1016/j.dr.2016.12.001

Spinelli, M., Fasolo, M., Tagini, A., Zampini, L., Suttora, C., Zanchi, P., & Salerni, N. (2016). Linguistic and prosodic aspects of child-directed speech: The role of maternal child-rearing experiences. *European Journal of Developmental Psychology*, *13*(2), 183–196.

Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, *10*(1), 1–15. https://doi.org/10.1017/S0305000900005092

Styles, S., & Plunkett, K. (2009). What is 'word understanding' for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report. *Journal of Child Language*, *36*(4), 895–908.

Suttora, C., Salerni, N., Zanchi, P., Zampini, L., Spinelli, M., & Fasolo, M. (2017). Relationships between structural and acoustic properties of maternal talk and children's early word recognition. *First Language*, *37*(6), 612–629.

Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2018). Routine language: Speech directed to infants during home activities. *Child Development*. doi:https://doi.org/10.1111/cdev.13089

Tang, J. S. Y., & Maidment, J. A. (1996). Prosodic aspects of child-directed speech in Cantonese. *University College London: Speech, hearing and language—Work* *in progress*, *9*, 257–276.

Tang, P., Xu Rattanasone, N., Yuen, I., & Demuth, K. (2017). Phonetic enhancement of Mandarin vowels and tones: Infant-directed speech and lombard speech. *The Journal of the Acoustical Society of America*, *142*(2), 493–503. https://doi.org/10.1121/1.4995998

Tardif, T., Fletcher, P., Liang, W., & Kaciroti, N. (2009). Early vocabulary development in Mandarin (Putonghua) and Cantonese. *Journal of Child Language*, *36*(5), 1115–1144.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5

Thorson, J. C. (2018). The role of prosody in early word learning: Behavioral evidence. In P. Prieto, & N. Esteve-Gibert (Ed.), *The development of prosody in first language acquisition* (Vol. 23, 1st ed., pp. 60–77). Amsterdam, The Netherlands: John Benjamins.

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review*, *9*(2), 335–340.

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, *11*(3), 188–195. https://doi.org/10.1111/1467-9280.00240

Vallotton, C. D., Decker, K. B., Kwon, A., Wang, W., & Chang, T. (2017). Quantity and quality of gestural input: Caregivers' sensitivity predicts caregiver-infant bidirectional communication through gestures. *Infancy*, *22*(1), 56–77. https://doi.org/10.1111/infa.12155

Van de Weijer, J. (1999). *Language input for word discovery* (Doctoral dissertation), Radboud University Nijmegen, Nijmegen. doi:10.17617/2.2057670.

Van der Hulst, H. (1984). *Syllable structure and stress in Dutch*. Dordrecht, The Netherlands: Foris.

Van Rooijen, R., Bekkers, E.L., & Junge, C.M.M. (2018, July). *Two-year-olds learn faster from their mothers in novel word learning situations.* Conference talk presented at the XXI biannual meeting of the International Congress of Infant Studies, Philadelphia, PA.

Venables, W. N. & Ripley, B. D. (2002) *Modern Applied Statistics with S*. Fourth edition. Berlin: Springer.

Wang, Y., Llanos, F., & Seidl, A. (2017). Infants adapt to speaking rate differences in word segmentation. *The Journal of the Acoustical Society of America*, *141*(4), 2569–2578.

Wang, Y., Seidl, A., & Cristia, A. (2016). Acoustic characteristics of infant-directed speech as a function of prosodic typology. In J. Heinz, R. Goedemans, & H. van de Hulst (Ed.). *Dimensions of phonological stress*. (pp. 311–324). Cambridge, MA: Cambridge University Press.

Weisleder, A., & Fernald, A. (2013). Talking to children matters early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152.

Werker, J. F. (2018). Perceptual beginnings to language acquisition. *Applied Psycholinguistics*, *39*(4), 703–728. https://doi.org/10.1017/S0142716418000152

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63.

Werker, J. F., Cohen, L. B., Lloyd, V., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology*, *34*(6), 1289–1309. https://doi.org/10.1037/0012-1649.34.6.1289

Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, *17*(3), 323–333.

Wong, P., & Ng, H.-Y. (2017). Do mothers enhance the tonal contrasts in their monosyllabic Cantonese tones directed to their infants? *The Journal of the Acoustical Society of America*, *141*(5), 3748–3748. doi:10.1121/1.4988265

Xu Rattanasone, N., Burnham, D., & Reilly, R. G. (2013). Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age. *Journal of Phonetics*, *41*(5), 332–343.

Xu, N. (2008). *Tones and vowels in Cantonese infant-directed speech: Hyperarticulation during the first 12 months of infancy* (Doctoral dissertation, University of Western Sydney).

Xu, N., Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2013). Vowel hyperarticulation in parrot-, dog- and infant-directed speech. *Anthrozoös*, *26*(3), 373–380. https://doi.org/10.2752/175303713X13697429463592

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*(1), 61–83.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics*, *27*(1), 55–105. https://doi.org/10.1006/jpho.1999.0086

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, *68*(2), 123–139.

Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.

Zangl, R., & Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6-and 13-month-old infants. *Infancy*, *11*(1), 31–62.

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, *6*(2), 179–208.

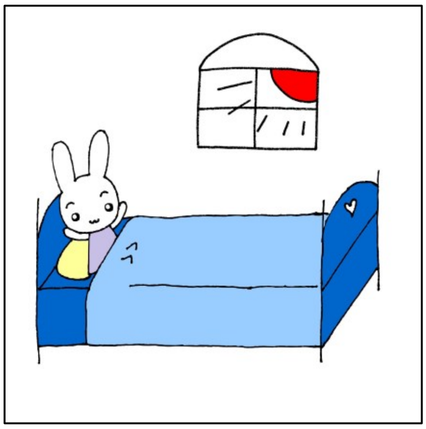
Zhang, J. (2014). *A sociophonetic study on tonal variation of the Wúxî and Shànghǎi dialects*. (Doctoral dissertation, Utrecht University). Retrieved from https://www.lotpublications.nl/Documents/370_fulltext.pdf.

Zink, I., & Lejaegere, M. (2002). *N-CDI's lijsten voor communicatieve ontwikkeling*. ("Dutch MacArthur CDI's for communicative development"). Leuven, Belgium: Acco.

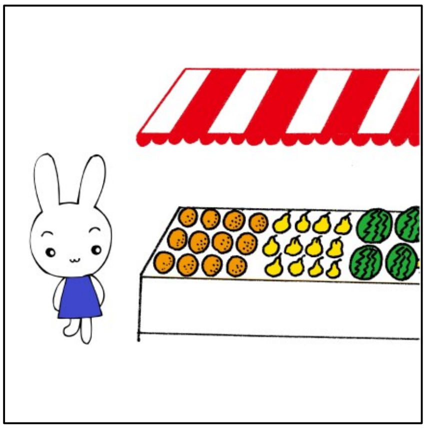
Appendix A

Picture books

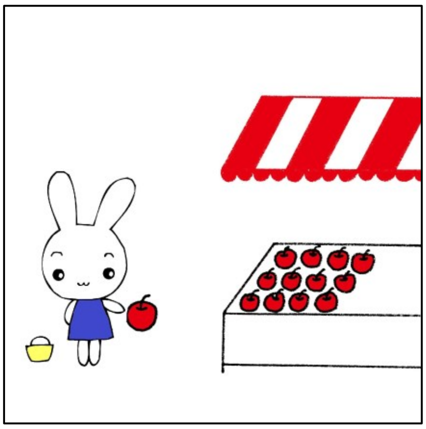
Book 1: *Konijntje heeft geluk* ("Bunny is lucky"). This book was used for Dutch 18-month-old children.



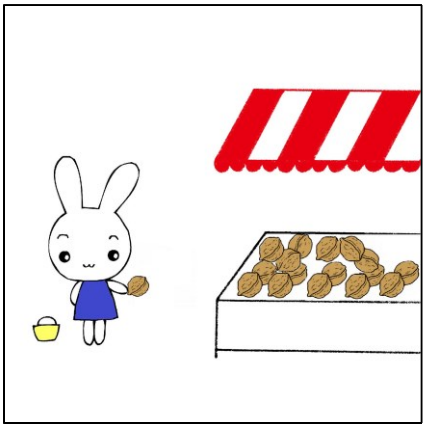
markt



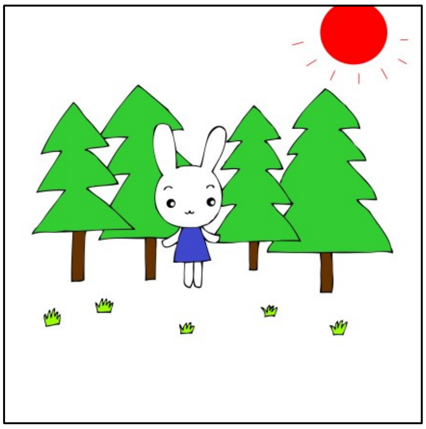
appel



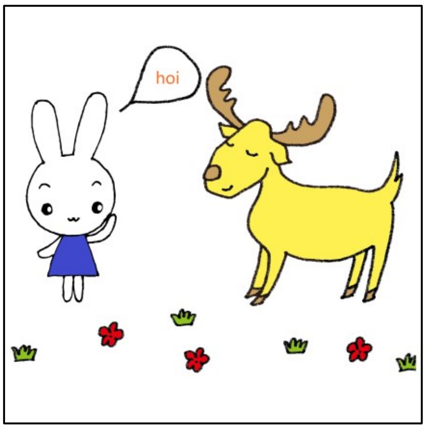
walnoot



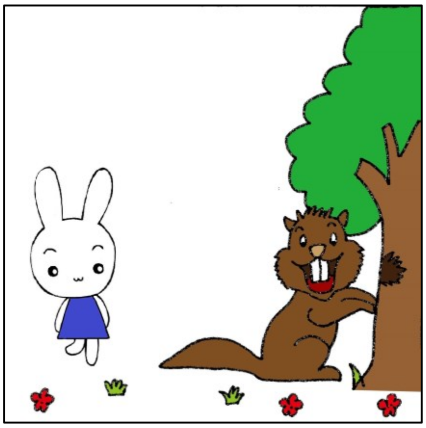
bos



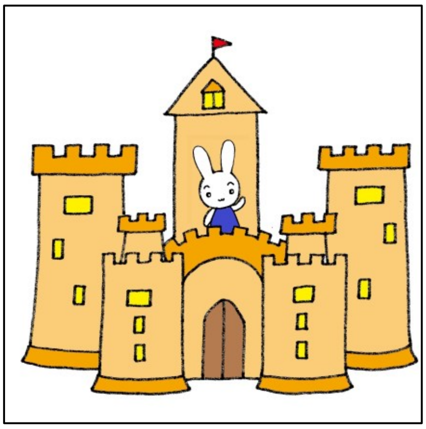
eland



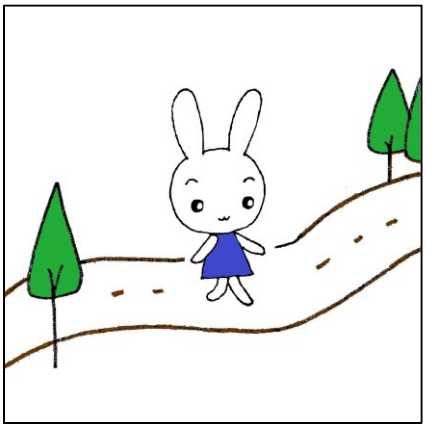
bever



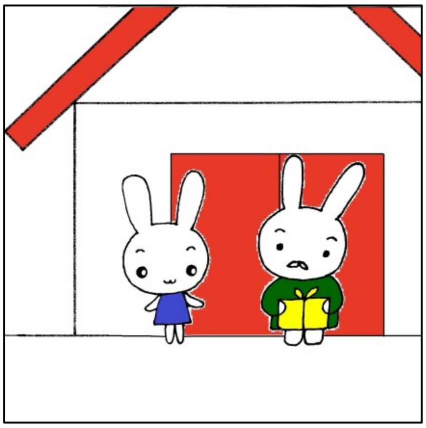
kasteel



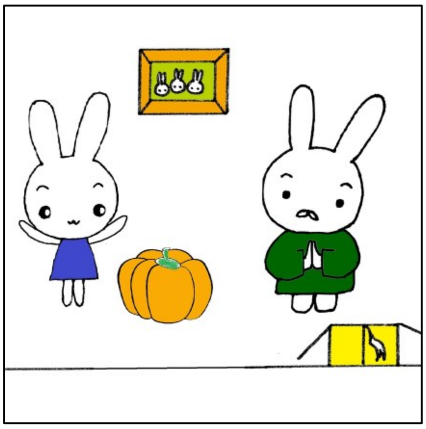
naar huis



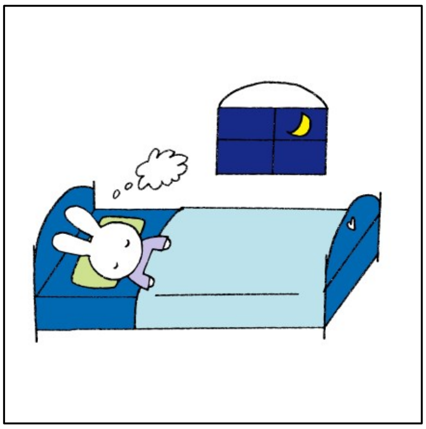
opa



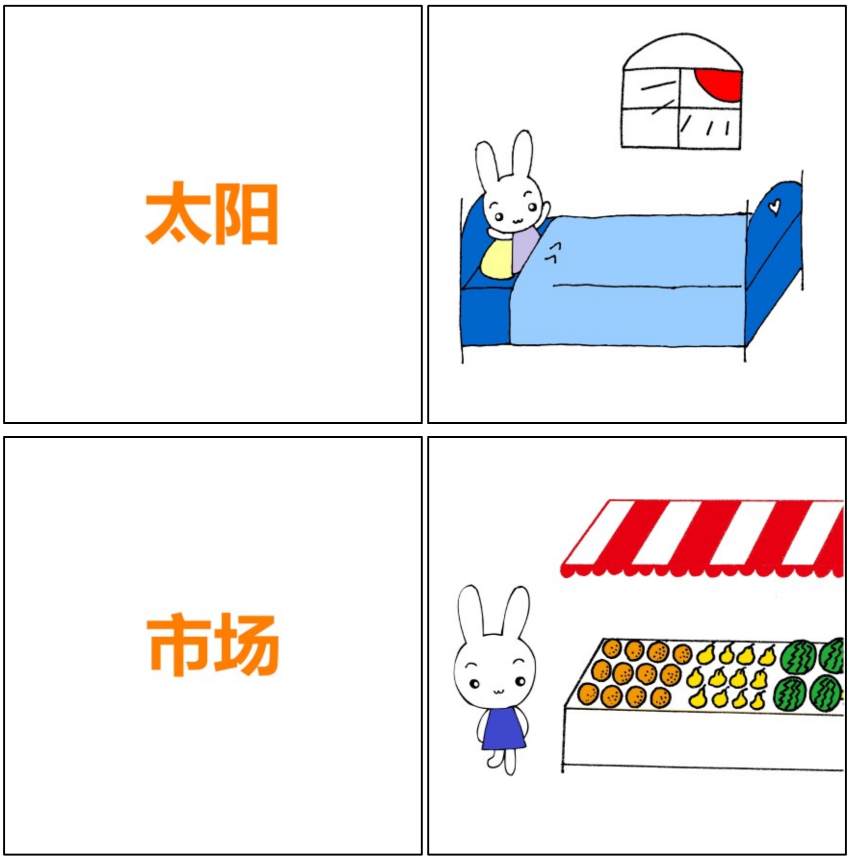
pompoen



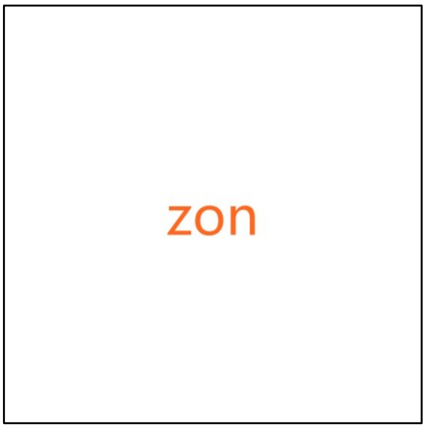
slapen

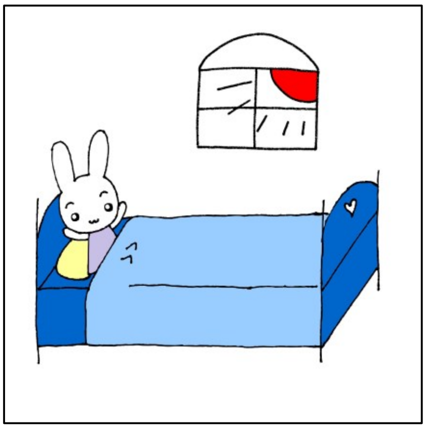


Example pages of Book 2: *小兔子的一天* ("Bunny's day"). This book is an adaption of Book 1 (Dutch) to Mandarin Chinese and it was used for Mandarin Chinese 18- and 24-month-old children.

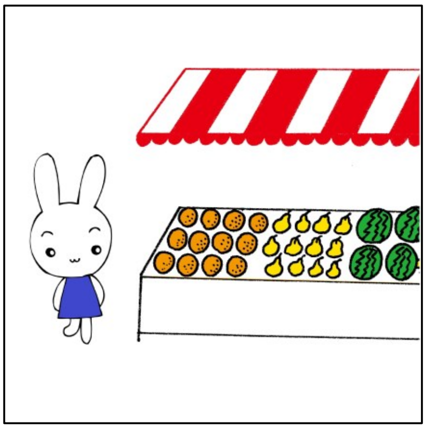


Book 3: *Konijntje heeft geluk* ("Bunny is lucky"). This book was used for Dutch 24-month-old children. The book structure and familiar target words were the same with Book 1 (for Dutch 18-month-old children), only unfamiliar target words were different.

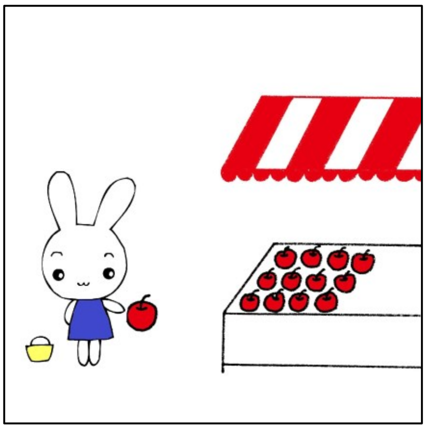




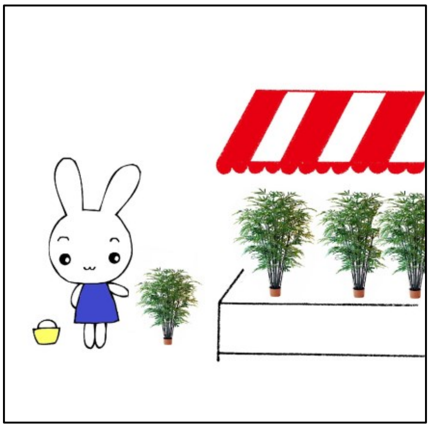
markt



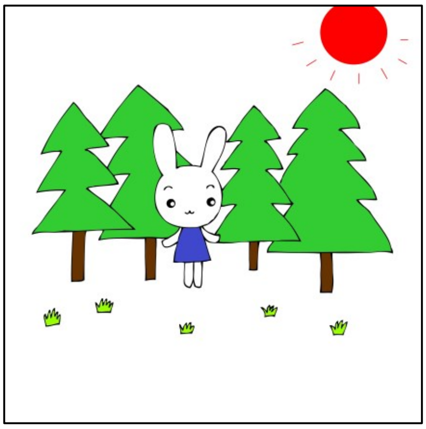
appel



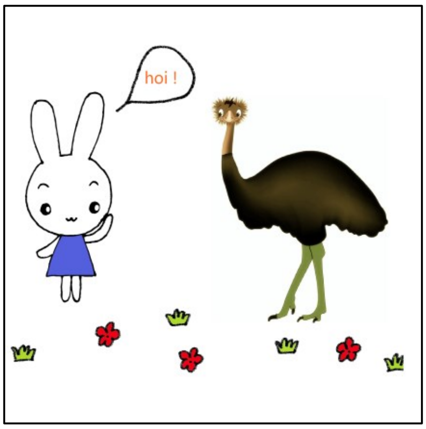
bamboe



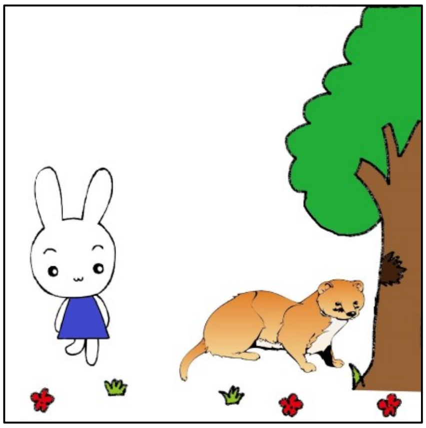
bos



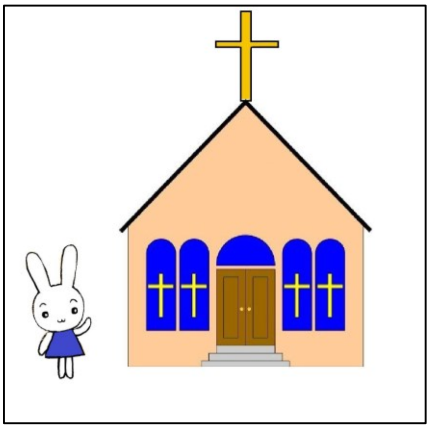
emoe



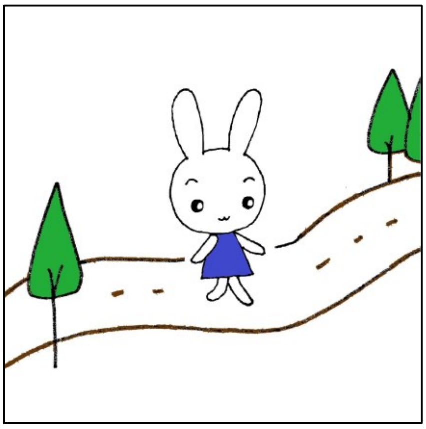
wezel



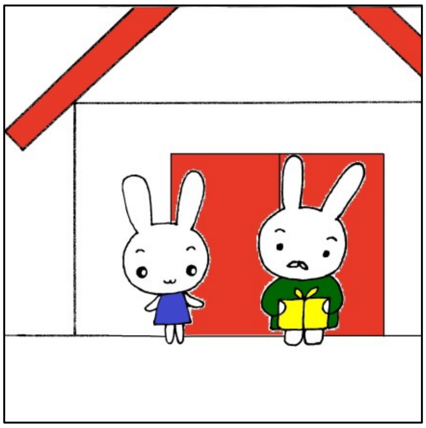
kapel



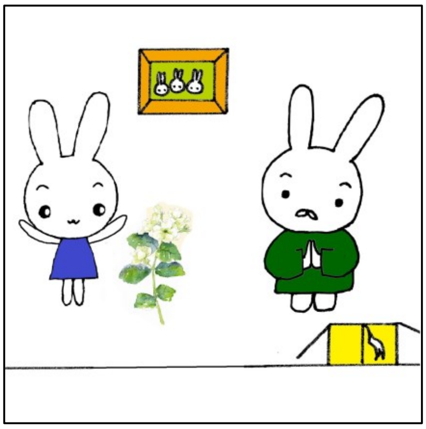
naar huis



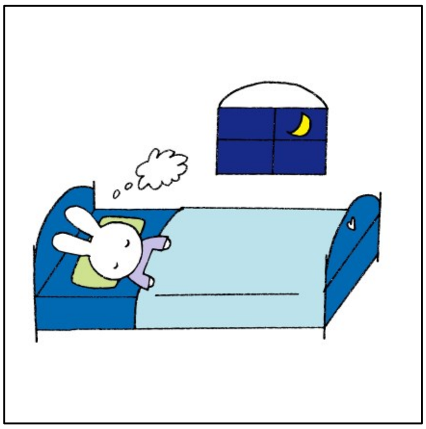
opa



jasmijn



slapen



Appendix B

Supplementary materials for Chapter 4

Supplementary Table 1. Model for Maximum F0 (Hz) for the 18-month-old group

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	16.809	0.390	43.056	<0.001***
Condition (IDS)	1.401	0.398	3.516	0.002**
Tone2	-0.116	0.342	-0.339	0.735
Tone3	-1.097	0.333	-3.293	0.001**
Tone4	1.017	0.319	3.191	0.002**

Supplementary Table 2. Model for Maximum F0 (Hz) for the 24-month-old group

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	17.420	0.365	47.710	<0.001***
Tone2	-0.872	0.388	-2.247	0.025*
Tone3	-0.727	0.380	-1.914	0.057
Tone4	1.227	0.375	3.273	0.001**

Supplementary Table 3. Model for Minimum F0 (Hz) for the 18-month-old group

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	16.080	0.325	49.530	<0.001***
Condition (IDS)	0.589	0.224	2.630	0.010**
Tone2	-1.907	0.313	-6.091	<0.001***
Tone3	-2.834	0.305	-9.283	<0.001***
Tone4	-0.834	0.292	-2.852	0.005*

Supplementary Table 4. Model for Minimum F0 (Hz) for the 24-month-old group

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	16.282	0.311	52.031	<0.001***
Tone2	-1.775	0.321	-5.525	<0.001***
Tone3	-2.844	0.315	-9.037	<0.001***
Tone4	-1.292	0.311	-4.161	<0.001***

Supplementary Table 5. Final model for Maximum F0 (ERB) for the 18-month-old and 24-month-old group

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	2.629	0.044	59.459	<0.001***
Condition (IDS)	-0.148	0.043	3.405	<0.002***
Tone2	-0.053	0.029	-1.865	0.063
Tone3	-0.100	0.028	-3.591	0.001**
Tone4	0.125	0.027	4.601	<0.001***
Age (24m)	0.069	0.060	1.157	0.255
Condition(IDS):Age(24m)	-0.160	0.063	-2.532	0.016*

Supplementary Table 6. Final model for Minimum F0 (ERB) for the 18-month-old and 24-month-old group

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	2.558	0.043	60.138	<0.001***
Condition (IDS)	0.067	0.030	2.219	0.030*
Tone2	-0.239	0.040	-6.159	<0.001***
Tone3	-0.365	0.036	-10.062	<0.001***
Tone4	-0.136	0.033	-4.122	<0.001***
Age (24m)	0.045	0.051	0.881	0.384
Condition(IDS):Age(24m)	-0.112	0.043	-2.607	0.012*

Supplementary Table 7. Model for Maximum F0 (ERB) for the 18-month-old group

Parameters	***Estimate***	***SE***	***t-value***	***p***
	Fixed factors			
(Intercept)	2.626	0.044	59.825	<0.001***
Condition (IDS)	0.148	0.043	3.407	0.003**
Tone2	-0.009	0.038	-0.233	0.816
Tone3	-0.120	0.037	-3.224	0.001**
Tone4	0.116	0.035	3.268	0.001**

Supplementary Table 8. Model for Maximum F0 (ERB) for the 24-month-old group

Parameters	*Estimate*	*SE*	*t-value*	*p*
	Fixed factors			
(Intercept)	2.694	0.040	67.234	<0.001***
Tone2	-0.100	0.043	-2.250	0.025*
Tone3	-0.085	0.042	-2.066	0.040*
Tone4	0.130	0.041	2.829	0.005**

Supplementary Table 9. Model for Minimum F0 (ERB) for the 18-month-old group

Parameters	*Estimate*	*SE*	*t-value*	*p*
	Fixed factors			
(Intercept)	2.537	0.039	64.388	<0.001***
Condition (IDS)	0.067	0.028	2.361	0.019*
Tone2	-0.224	0.041	-5.507	<0.001***
Tone3	-0.336	0.040	-8.497	<0.001***
Tone4	-0.089	0.038	-2.352	0.019*

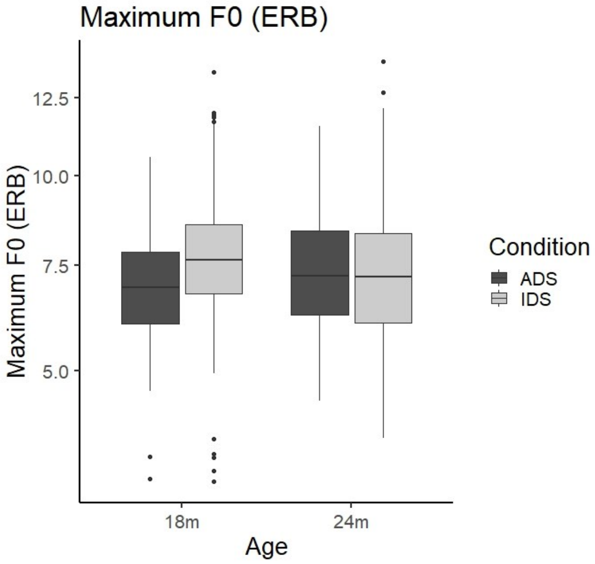
Supplementary Table 10. Model for Minimum F0 (ERB) for the 24-month-old group

Parameters	*Estimate*	*SE*	*t-value*	*p*
(Intercept)	2.566	0.038	67.497	<0.001***
Tone2	-0.214	0.040	-5.378	<0.001***
Tone3	-0.354	0.039	-9.097	<0.001***
Tone4	-0.158	0.038	-4.109	<0.001***

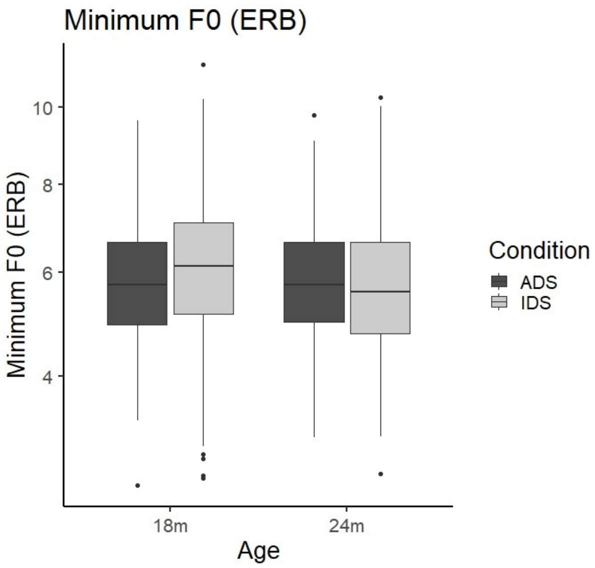
Supplementary Table 11. Final model for F0 range (ERB) for the 18-month-old and 24-month-old group

Parameters	*Estimate*	*SE*	*t-value*	*p*
	Fixed factors			
(Intercept)	1.130	0.740	15.293	<0.001***
Condition (IDS)	0.107	0.046	2.325	0.020*
Tone3	0.036	0.082	0.441	0.660
Tone4	0.019	0.078	0.252	0.801
Age (24m)	-0.169	0.098	-1.726	0.087
Tone3:Age (24m)	0.246	0.115	2.139	0.033*
Tone4:Age (24m)	0.338	0.112	3.043	0.002**

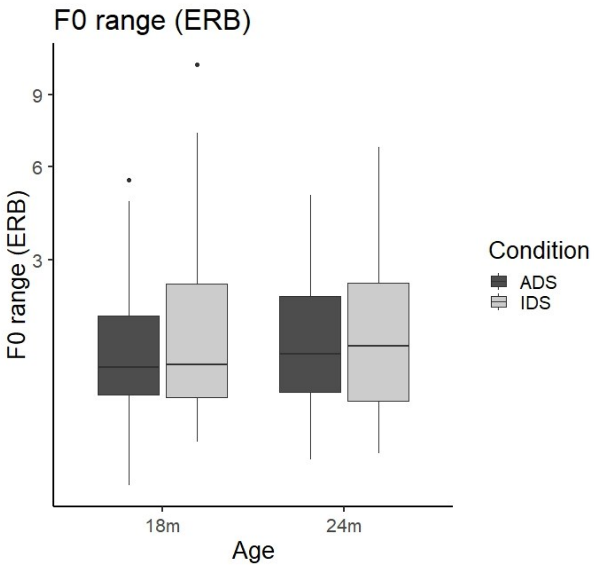
Supplementary Figure 1. Box plots of Maximum F0 (ERB) for ADS and IDS addressing 18-month-old and 24-month-old children



Supplementary Figure 2. Box plots of Minimum F0 (ERB) for ADS and IDS addressing 18-month-old and 24-month-old children.



Supplementary Figure 3. Box plots of F0 range (ERB) for ADS and IDS addressing 18-month-old and 24-month-old children**.**



Nederlandse Samenvatting

(Summary in Dutch)

De rol van prosodische input bij het leren van woorden: Een cross-linguïstisch onderzoek naar kindgerichte spraak in het Nederlands en het Mandarijn

Deze dissertatie onderzoekt de rol van prosodische input bij het leren van woorden door kinderen, met een focus op de prosodie van kindgerichte spraak (KGS) in woordleersituaties, waarin moeders voor hun kind onbekende woorden introduceren. Een belangrijk aspect van dit onderzoek is de cross-linguïstische aanpak waarbij KGS in twee typologisch sterk verschillende talen wordt onderzocht: het Nederlands en het Mandarijn Chinees.

Voor het leren van woorden moeten kinderen nieuwe woordlabels uit het taalaanbod verbinden aan begrippen, een proces dat "word-to-world mapping" wordt genoemd (Brooks & Kempe, 2014, p. 679). Naast *welke* woordlabels de kinderen te horen krijgen, zou het voor het leren ook nog uit kunnen maken *hoe* die nieuwe woorden zijn uitgesproken. Wanneer ze tegen hun kinderen praten gebruiken moeders een speciale stijl: "kindgerichte spraak" (KGS), soms ook wel "babypraat" genoemd. Deze spreekstijl onderscheidt zich op een aantal punten van volwassenegerichte spraak (VGS). KGS heeft in vergelijking met VGS onder andere meer verkleinwoordjes, overgearticuleerde klinkers, kortere uitingen, meer herhaling, en een lagere grammaticale complexiteit. De opvallendste eigenschap van KGS is overdreven prosodie, waaronder een hogere toonhoogte, een wijder toonhoogtebereik, en een langzamere spreeksnelheid dan bij VGS (zie de reviews in Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015; Soderstrom, 2007).

Tot nu toe is de algemene overdreven prosodie van KGS gevonden in bijna alle onderzochte talen, met één uitzondering, namelijk het Quiché Mayaans (Ingram, 1995; Bernstein Ratner & Pye, 1984). Onderzoekers hebben drie functies van KGS voorgesteld: het trekken van de aandacht van het kind, het overbrengen van positief affect, en het ondersteunen van taalverwerving. De rol van prosodische input bij het vroege woorden leren is echter nog niet volledig doorgrond. Specifieker gezegd zijn er op dit punt nog vijf belangrijke gaten in de onderzoeksliteratuur. Ten eerste heeft tot nu toe geen enkele studie de prosodie van KGS specifiek onderzocht in woordleersituaties, waarin moeders voor hun kind onbekende woorden introduceren. Ten tweede richten studies naar de prosodie van KGS zich vaak op een enkele taal (meestal Amerikaans Engels) zonder cross-linguïstisch onderzoek onder typologisch van elkaar verschillende talen. Ten derde richten veel KGS-studies zich op het eerste levensjaar, waardoor er relatief weinig onderzoek is gedaan naar KGS die wordt gesproken tegen kinderen die in de "vocabulairespurt" zitten die loopt van 18 tot 24 maanden oud. Ten vierde: prototypische KGS helpt bij het online woord-naar-object mappen van Amerikaans Engelssprekende kinderen (e.g., Graf Estes & Hurley, 2013; Ma et al., 2011), maar het is niet bekend of er vergelijkbare effecten optreden voor andere talen. Ten slotte is het tot nu toe niet bekend, ondanks sterke evidentie dat de kwantiteit van de input (d.w.z. het aantal woorden) gecorreleerd is met de grootte van het vocabulaire van het kind, of er correlaties bestaan tussen de prosodie van KGS en de taaluitkomsten van kinderen, waaronder grootte van het vocabulaire en vermogen om online woorden naar objecten te mappen.

Het overkoepelende doel van deze dissertatie is om te bestuderen wat de rol van prosodische input is bij het woorden leren door kinderen. Een bijkomend doel is om ons begrip te verbeteren van de universele en taalspecifieke kanten van KGS-prosodie. Om deze doelen te behalen en de hierboven genoemde gaten in de literatuur te vullen, hebben we een cross-linguistïsch onderzoek uitgevoerd naar KGS in twee typologisch verschillende talen: Nederlands (een Germaanse taal met lexicale klemtoon en klemtoongebaseerd spraakritme) en Mandarijn Chinees (een Sinitische taal met lexicale toon en een lettergreepgebaseerd spraakritme).

Dit onderzoek omvat drie hoofdexperimenten waarover wordt gerapporteerd in vijf empirische studies (Hoofdstukken 2 t/m 6). De drie experimenten en vijf studies worden hieronder samengevat.

Experimenten

Voor dit onderzoek zijn drie experimenten uitgevoerd. In Experiment 1 namen Nederlandssprekende dyades deel aan een semi-spontane boekvoorleestaak op de tijden dat het kind 18 en 24 maanden oud was (longitudinaal ontwerp). Om een woordleersituatie op te zetten bevatte het verhaal in het boek voor de kinderen zowel bekende als onbekende woorden. De resultaten van Experiment 1 worden gerapporteerd in Hoofdstukken 2, 3, 5 en 6, die focussen op verschillende prosodische eigenschappen van de spraakdata.

In Experiment 2 namen Mandarijn Chineessprekende moeders deel aan een semi-spontane boekvoorleestaak. De materialen en spraakelicitatiemethoden waren vergelijkbaar met die van Experiment 1. Het grote verschil tussen Experiment 1 en Experiment 2 is dat Experiment 2 een cross-sectioneel ontwerp had in plaats van een longitudinaal ontwerp. De resultaten van Experiment 2 worden gerapporteerd in Hoofdstukken 2, 3, 4 en 5, die focussen op verschillende aspecten van de spraakdata.

Experiment 3 is een woordleerexperiment dat gebruikt maakt van het Intermodal Preferential Looking Paradigm, waarin de woord-naar-object-mapping-prestaties van kinderen werden vergeleken voor VGS en KGS-condities. De deelnemers aan dit experiment zijn een subset van de deelnemers van Experiment 1. Deze deelnemers namen deel aan Experiment 1 tijdens zowel de leeftijd van 18 maanden als van 24 maanden van het kind, en werden daarnaast getest op hun woordleerprestaties toen de kinderen 24 maanden oud waren. De resultaten van Experiment 3 worden gerapporteerd in Hoofdstuk 6.

In addition, we collected vocabulary information from all participants in the three experiments. All mothers filled in Dutch (N-CDI; Zink & Lejaegere, 2002) or Mandarin Chinese (M-CDI; Tardif, Fletcher, Liang, & Kaciroti, 2009) versions of MacArthur-Bates Communicative Development Inventories (CDI; Fenson, Bates, Dale, Marchman, & Reznick, 2007).

Verder hebben we informatie verzameld over de vocabulairekennis van alle deelnemers in alledrie de experimenten. Alle moeders beantwoordden de Nederlandse (N-CDI; Zink & Lejaegere, 2002) of Mandarijn Chinese versie (M-CDI; Tardif, Fletcher, Liang, & Kaciroti, 2009) van de MacArthur-Bates Communicative Development Inventories (CDI; Fenson, Bates, Dale, Marchman, & Reznick, 2007).

Empirische studies

Hoofdstuk 2

Hoofdstuk 2 heeft als doel om de spreeksnelheid van KGS in woordleersituaties te onderzoeken. Specifiek gezegd zijn we uitgegaan van de vragen (1) of Nederlandse en Chinese moeders specifiek langzamer praten om onbekende woorden te onderstrepen, (2) of KGS langzamer gesproken wordt dan VGS in het Nederlands en het Mandarijn Chinees, (3) of de spreeksnelheid van KGS en KGS-in-woordleersituaties verschilt voor moeders wanneer ze tegen 18 of 24 maanden oude kinderen praten.

De resultaten in Hoofdstuk 2 wijzen uit dat, over het geheel genomen, Nederlandse KGS gericht tot 18 en 24 maanden oude kinderen langzamer wordt gesproken vergeleken met VGS, maar er waren geen significante verschillen in spreeksnelheid tussen Mandarijn Chinese VGS en KGS voor de beide leeftijdsgroepen. In woordleersituaties vertraagden Nederlandse moeders specifiek wanneer ze onbekende woorden introduceerden in KGS, vergeleken met bekende woorden, wat kinderen zou kunnen helpen bij het leren van woorden. Daarentegen vonden we geen evidentie dat Mandarijn Chinese KGS vertraagt om onbekende woorden te onderstrepen. Voor zowel Nederlands als Mandarijn Chinees bieden onze resultaten geen evidentie van leeftijdsgerelateerde veranderingen in de spreeksnelheid van KGS gericht aan 18 en 24 maanden oude kinderen. Deze resultaten suggereren dat temporale aanpassingen in KGS, met name het effect van bekendheid van het woord op de spreeksnelheid van KGS, kunnen verschillen tussen talen.

Hoofdstuk 3

Waar Hoofdstuk 2 zich richtte op het temporale aspect van KGS, onderzoekt Hoofdstuk 3 de toonhoogte-eigenschappen ervan. In het bijzonder behandelden we de vragen (1) of moeders verschillend omgaan met toonhoogte bij de uitspraak van bekende en onbekende woorden in KGS vergeleken met VGS, (2) hoe Mandarijn Chinese en Nederlandse KGS verschillen in hun toonhoogte-eigenschappen in woordleersituaties waarin moeders onbekende woorden introduceren aan hun kinderen, en (3) of toonhoogte-eigenschappen van KGS en KGS-in-woordleersituaties veranderen wanneer moeders praten met 24 maanden oude kinderen ten opzichte van 18 maanden oude kinderen.

De resultaten wijzen uit dat, over het geheel bekeken, de toonhoogte-overdrijving van KGS verschillend was tussen Nederlands en Mandarijn Chinees: Nederlandse KGS gericht aan 18 en 24 maanden oude kinderen had een hogere gemiddelde toonhoogte dan VGS, ook al waren de verschillen in toonhoogte tussen de twee spreekstijlen kleiner bij 24 maanden oude kinderen dan bij 18 maanden oude kinderen. Er waren geen significante verschillen qua toonhoogtebereik tussen de twee leeftijdsgroepen. Mandarijn Chinese KGS gesproken tegen 18 maanden oude kinderen had een hoger toonhoogtegemiddelde en een wijder toonhoogtebereik vergeleken met VGS, maar Mandarijn Chinese KGS gesproken tegen 24 maanden oude kinderen was vergelijkbaar met VGS wat betreft zowel toonhoogtegemiddelde als –bereik. De prosodie in woordleersituaties verschilde ook tussen Nederlands en Mandarijn Chinees. Nederlandse moeders hadden een lager toonhoogtegemiddelde specifiek voor onbekende woorden bij zowel 18 als 24 maanden oude kinderen. De Mandarijn Chinese moeders, daarentegen, hadden een hoger toonhoogtegemiddelde specifiek wanneer ze onbekende woorden introduceerden aan 18 maanden oude kinderen, en ze hadden een wijder toonhoogtebereik wanneer ze onbekende woorden introduceerden aan 24 maanden oude kinderen. Deze resultaten suggereren dat de toonhoogte-eigenschappen van KGS specifiek voor woordleersituaties verschillen tussen Nederlands en Mandarijn Chinees.

De resultaten van Hoofdstuk 2 en Hoofdstuk 3 wijzen er samen op dat Nederlandse en Chinese moeders de vocabulairekennis van hun kinderen aanvoelen en hun prosodie daarop afstemmen in woordleersituaties. KGS-prosodie in woordleersituaties komt echter verschillend tot uiting in de verschillende talen. Op basis van deze bevindingen is er behoefte aan verder onderzoek naar de vraag of KGS-prosodie een taalkundige functie heeft en naar de mate waarin prosodische cues een ondersteunende rol kunnen spelen bij het leren van woorden.

Hoofdstuk 4

Mandarijn Chinees is een toontaal, waarin lexicale toon een cruciale rol speelt om de betekenis van woorden te bepalen. Daarom is het belangrijk om onderzoek te doen naar lexicale tonen in KGS in toontalen. Hoofdstuk 4 rapporteert over een studie naar lexicale tonen in Mandarijn Chinese KGS gesproken tegen 18 en 24 maanden oude kinderen.

De resultaten van Hoofdstuk 4 wezen uit dat lexicale tonen in KGS gericht aan 18 maanden oude kinderen een hogere minimum F0 hadden, alsook een hogere maximum F0 en een wijder toonhoogtebereik, vergeleken met lexicale tonen in VGS. Lexicale tonen in KGS gesproken tegen 24 maanden oude kinderen waren overeenkomstiger met KGS-tonen wat betreft toonhoogtegrenzen: er waren geen verschillen in minimum F0 of maximum F0 tussen KGS en VGS. Het toonhoogtebereik was echter nog steeds wijder. Deze observaties suggereren dat lexicale tonen in Mandarijn Chinese KGS overgearticuleerd zijn voor zowel de 18 maanden oude als de 24 maanden oude kinderen, ondanks de verschillen in het toonhoogteniveau over tijd. Mandarijn Chinese moeders overarticuleren lexicale tonen in KGS wanneer ze tegen peuters praten en ondersteunen daarmee mogelijk toonverwerving en het leren van woorden.

Hoofdstuk 5

Hoewel het is aangetoond dat de prosodische features van prototypische KGS de aandacht trekken van baby's en positieve affect overbrengen, wordt er nog steeds gedebatteerd over de vraag of KGS-prosodie specificiek taalkundige functies heeft. In het bijzonder is het nog onduidelijk of er correlaties bestaan tussen de prosodie van KGS en de vocabulairegrootte van het kind. In Hoofdstuk 5 onderzochten we voor zowel het Nederlands als het Mandarijn Chinees de correlaties tussen KGS-prosodie en de vocabulairegrootte van de kinderen. De voorspellers waren prosodische hyperscores. Een hyperscore werd berekend door voor elke moeder haar KGS-score (b.v. gemiddelde toonhoogte) te delen door haar VGS-score. Specifiek hebben we twee hyperscores berekend: de algemene hyperscore (die algemene prosodische overdrijving weergeeft) en de onbekendheidshyperscore (die prosodische overdrijving in woordleersituaties weergeeft). Eerst hebben we gekeken of er concurrente correlaties bestaan tussen de prosodische hyperscores en de vocabulairegrootte van de kinderen. Vervolgens hebben we onderzocht of de prosodische hyperscores wanneer de kinderen 18 maanden oud zijn kunnen voorspellen wat de groei van het vocabulaire is tussen 18 en 24 maanden oud.

De Nederlandse resultaten lieten geen significante correlaties zien tussen de algemene hyperscores en de grootte van het kindervocabulaire, noch als concurrente correlatie, noch longitudinaal. Echter waren er wel significante correlaties tussen de Nederlandse onbekendheidshyperscore en de grootte van het kindervocabulaire. Niet alleen waren de Nederlandse onbekendheidshyperscores significant concurrent gecorreleerd met de grootte van het vocabulaire van de kinderen op 18 en 24 maanden oud, ook waren deze scores significant voorspellend voor de groei van het kindervocabulaire tussen 18 en 24 maanden oud. Voor Mandarijn Chinees bevatten onze observaties geen significante correlaties tussen de algemene hyperscores en de grootte van het vocabulaire van de kinderen, en ook waren er geen significante correlaties tussen de onbekendheidshyperscores en de vocabulairegrootte van de kinderen. Samengevat suggereren deze resultaten dat de prosodie van (Nederlandse) KGS specifiek voor woordleersituaties taalkundige doelen zou kunnen dienen.

Hoofdstuk 6

Hoofdstuk 6 deed verslag van een woordleerexperiment gericht op de vraag of prototypische Nederlandse KGS de woord-naar-object-mapping van 24 maanden oude kinderen ondersteunt in vergelijking met VGS. Onze bevindingen wijzen uit dat Nederlandse 24 maanden oude kinderen consistent nieuwe woord-naar-object-mappings leren in zowel de VGS- en de KGS-conditie. Er was slechts een klein ondersteunend effect van KGS op één type meting (gemiddelde “single longest look”), maar over het algemeen boden onze resultaten geen sterke evidentie voor de claim dat prototypische KGS het leren van woorden ondersteunt. Op basis van de hyperscores uit Hoofdstuk 5 en de resultaten over het leren van woorden van Experiment 3 hebben we verder verkend of woordleerprestaties in VGS- en KGS-condities gerelateerd zijn aan het taalaanbod dat kinderen ontvangen hebben. Specifiek hebben we gekeken of de mate van prosodische overdrijving in KGS (weergegeven door de algemene hyperscore) en de mate van prosodische overdrijving in KGS in woordleersituaties (weergegeven door de onbekendheidshyperscore) significante voorspellers zijn van de prestaties van kinderen bij woord-naar-object-mapping in VGS en KGS. De resultaten lieten geen significante correlaties zien tussen de algemene hyperscores of de onbekendheidshyperscores en de woordleerprestaties van de kinderen in KGS of VGS, tenminste niet voor dit kleine sample. De bevindingen suggereren dat zelfs wanneer kinderen goed in staat zijn nieuwe woorden te leren uit VGS, er nog steeds een ondersteunend effect zou kunnen bestaan van prototypische KGS op het online leren van woorden door kinderen. Of deze bevindingen ook gelden voor andere talen of leeftijdsgroepen behoeft echter nog verder onderzoek.

Conclusie

Het hoofddoel van het huidige werk was om de rol van prosodische input te onderzoeken bij het leren van woorden door kinderen, door een onderzoek naar Nederlandse en Mandarijn Chinese KGS. Deze dissertatie vult een aantal gaten in de wetenschappelijke literatuur en voegt een aantal zaken toe.

Ten eerste biedt de dissertatie inzichten in de eigenschappen en mogelijke functies van KGS in de taalverwerving van kinderen door de aard te onthullen van prosodische input in woordleersituaties, door voort te borduren op eerdere literatuur over het effect van prototypische KGS op online woorden leren, en door voor het eerst significante correlaties aan te tonen tussen KGS-prosodie in woordleersituaties en de grootte van het vocabulaire van kinderen. Hoewel eerdere studies uit de literatuur suggereerden dat de algemene overdreven prosodie van KGS vooral samenhangt met aandacht en affect, suggereren onze bevindingen dat prosodische input specifiek in woordleersituaties gerelateerd zou kunnen zijn aan taalkundige functies.

Ten tweede bevat deze dissertatie consequenties over de universaliteit en taalspecifiekheid van KGS. Hoewel we evidentie vinden voor algemene prosodische overdrijving in zowel het Nederlands als het Mandarijn Chinees, wat in overeenstemming is met de universaliteit van KGS-prosodie, bieden onze bevindingen ook evidentie voor de taalspecifiekheid van KGS-prosodie. Cross-linguïstische verschillen in KGS-prosodie beperken zich niet tot prosodische overdrijving. Zo verschilden de specifieke prosodische cues die werden overdreven tussen Nederlandse en Mandarijn Chinese KGS, verschilden de manifestaties van leeftijdsgebonden veranderingen in KGS-prosodie in het Nederlands en Mandarijn Chinees, en verschilt met name de prosodie in woordleersituaties tussen deze twee talen. Gezien deze uitkomsten dient de taalspecifiekheid van KGS verder te worden onderzocht voor een grotere verzameling talen, met inachtneming van typologische verschillen in prosodie. Ook behoeft de vraag of prototypische KGS in alle talen het leren van woorden ondersteunt verder onderzoek, in meer talen.

Ten slotte heeft deze dissertatie zich gericht op de vocabulairespurt van kinderen in de periode van 18 tot 24 maanden oud. Zodoende draagt de dissertatie bij aan ons begrip over of en hoe KGS zich aanpast aan verschillende stadia van taalontwikkeling. De bevindingen suggereren dat de prosodie van KGS meer gaat lijken op die van VGS tijdens de groei van het kindervocabulaire in deze periode, en dat de kwaliteit van

KGS-prosodie een belangrijke rol speelt bij de groeispurt van het kindervocabulaire.

Aldus scheppen de bevindingen van deze dissertatie meer helderheid over de rol van prosodische input bij taalverwerving: de algemene overdreven prosodie van KGS is wellicht vooral nuttig voor het trekken van de aandacht van het kind, maar de prosodische input in woordleersituaties is fijntjes afgesteld op het dienen van een taalkundig doel en speelt mogelijk een grote rol bij het vroege woorden leren van kinderen.

Curriculum Vitae

Mengru Han was born on 10 July 1989 in Yichang, China. She studied Teaching Chinese as a Second Language at Nanjing University, China from 2007 to 2011. In 2010, she spent a semester at National University of Singapore as an exchange student under the Temasek Foundation International Leadership Enrichment and Regional networking (LEaRN) Programme. She moved to the Netherlands in 2011 and obtained a Research Master's degree in Linguistics from the Utrecht Institute of Linguistics (OTS), Utrecht University. She started her PhD project at Utrecht University in November 2013. Her PhD project is entitled "The role of prosodic input in word-learning" and this dissertation is the result of this research.