Language Learning and Development

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hlld20

Language Specificity of Infant-directed Speech: Speaking Rate and Word Position in Word-learning Contexts

Mengru Han, Nivja H. de Jong & René Kager

To cite this article: Mengru Han, Nivja H. de Jong & René Kager (2021) Language Specificity of Infant-directed Speech: Speaking Rate and Word Position in Word-learning Contexts, Language Learning and Development, 17:3, 221-240, DOI: 10.1080/15475441.2020.1855182

To link to this article: https://doi.org/10.1080/15475441.2020.1855182

LANGUAGE LEARNING AND DEVELOPMENT
2021, VOL. 17, NO. 3, 221–240
https://doi.org/10.1080/15475441.2020.1855182

Language Specificity of Infant-directed Speech: Speaking Rate and Word Position in Word-learning Contexts

Mengru Han[a,b,c], Nivja H. de Jong[d,e], and René Kager[b]

[a]Department of Chinese Language and Literature, East China Normal University, Shanghai, China; [b]Utrecht Institute of Linguistics (OTS), Utrecht University, Utrecht, The Nethelands; [c]Language, Cognition, and Evolution Lab, East China Normal University, Shanghai, China; [d]Leiden University Center for Linguistics (LUCL), Leiden University, Leiden, The Netherlands; [e]Leiden University Graduate School of Teaching (ICLON), Leiden University, Leiden, The Netherlands

ABSTRACT

Previous research indicates that infant-directed speech (IDS) is usually slower than adult-directed speech (ADS) and mothers prefer placing a focused word in isolation or utterance-final position in (English) IDS, which may benefit word learning. This study investigated the speaking rate and word position of IDS in two typologically-distinct languages: Dutch and Mandarin Chinese. We used a storybook-telling task to elicit semi-spontaneous ADS and IDS. The storybook contained target words that were familiar or unfamiliar to children. For each language, we asked: (1) whether IDS was slower than ADS; (2) whether mothers slowed down to highlight unfamiliar words; (3) whether the speaking rate of IDS changed between 18 and 24 months; and (4) whether mothers preferred placing unfamiliar words in isolation or utterance-final position in IDS. Results suggest that Dutch IDS, but not Mandarin Chinese IDS, was slower than ADS. Also, only Dutch mothers slowed down specifically when introducing unfamiliar words in IDS. In both languages, mothers placed target words in isolation (but not in utterance-final position) more frequently in IDS. These results suggest that the temporal modifications in IDS may vary across languages. Thus, language-specificity of IDS and its implications for language acquisition should be considered in future research.

Introduction

Infant-directed speech (IDS) has distinctive features compared to adult-directed speech (ADS). Prototypical IDS is characterized by an exaggerated prosody including pitch modifications (e.g., a higher pitch and a larger pitch range than IDS) and temporal modifications (e.g., a slower speaking rate) (see reviews in Golinkoff et al., 2015; Soderstrom, 2007). In addition to these prosodic modifications, mothers tend to place a contextually new word in isolation or utterance-final position in American English IDS (Fernald & Mazzie, 1991). The current study focused on the temporal modifications and word position of IDS. These features are presumed to facilitate children's word learning (Song et al., 2010; Zangl & Mills, 2007), however, no study has investigated them in a word-learning context, in which mothers introduce unfamiliar words to children.

This study aims at investigating the speaking rate of IDS specific to word-learning contexts. In addition, we explored whether mothers preferred putting unfamiliar words in isolation or utterance-final position in IDS. As existing studies on the speaking rate and word position of IDS often focus on

CONTACT René Kager r.w.j.kager@uu.nl Utrecht Institute of Linguistics (OTS), Utrecht University, Trans 10, Utrecht, 3512 JK, The Netherlands.

Portions of the data in this manuscript were published as: Han, M., de Jong, N. H., & Kager, R. (2018). 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.

a single language (mostly American English) while diverging by adopting different speech elicitation methods and measures, it is difficult to compare results directly from a cross-linguistic point of view. As such, we conducted a cross-linguistic study on Dutch and Mandarin Chinese IDS using similar speech elicitation methods.

Speaking rate of IDS across languages and ages

Even though it is widely assumed that IDS is slower than ADS, recent evidence suggests that the speaking rate of IDS is not consistently slower than ADS across languages, ages or across the board. 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 had shorter utterances, longer pauses, and a slower articulation rate (*excluding* silent pauses and measured in syllables per second) compared with ADS. American English IDS 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 toward 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 et al., 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 et al., 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 speaking rate of IDS may show age-related changes. In general, the prosodic properties of IDS including pitch modifications become more similar to ADS as children grow older and gain better linguistic competence (e.g., Han et al., 2020; Kitamura et al., 2002). So far only a few studies have explored the 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 IDS speaking rates in Korean, Tagalog, and Sri Lankan Tamil all tended to increase with age, with some individual variation among the small number of participants. Importantly, IDS speaking rates at 16 months were similar to ADS speaking rates in all three languages. As infants' phonetic perception shifts from language-universal to language-specific in the first year of life, known as "perceptual reorganization" (Werker & Tees, 1984), 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 phonetic categorization. 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.

Whether IDS is slower than ADS seems to be related to cross-linguistic differences or age-related changes, yet the mixed results reviewed above could be due to 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 et al., 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, ages, and across entire utterances based on previous studies. Crucially, cross-linguistic and longitudinal studies using similar speech elicitation methods are lacking regarding the speaking rate of IDS. Thus, our study sets out to fill this gap by examining IDS articulation 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 et al., 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 & 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.

Since we focused on 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 and examined the whether the speaking rate of IDS changed between the two ages.

The speaking rate and word position of IDS in word-learning contexts

It has long been suggested that IDS provides optimal information for children to learn languages. Regarding word learning, Fernald (2000) supported the view that caregivers modify their speech to potentially facilitate children's word recognition. Specifically, the author proposed that "[. . .] the idea that adults intuitively and dynamically accommodate speech to infants to make their meanings more accessible to inexperienced listeners is still plausible" (Fernald, 2000, p. 250). Following this proposal, we asked whether mothers vary speaking rate and word position in IDS to potentially facilitate word learning. Specifically, we examined these features in a word-learning context, in which mothers introduce unfamiliar words to children.

Previous studies have shown that a slower speaking rate of IDS facilitates children's word recognition and there is a correlation between the speaking rate of IDS and children's vocabulary size. 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 et al., 2005). Finally, a longitudinal study showed that slow speaking rate in IDS at 7 months predicts larger expressive vocabulary at 2 years of age (Raneri et al., 2020).

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 to promote 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 addition to temporal modifications, another strategy mothers may use is to vary the word position in IDS to support word learning. Specifically, words in isolation and in utterance-final position seem to benefit children's word learning. First, words in isolation are easier to segment compared to words embedded in continuous speech, as their word boundaries are marked by silent pauses before and after them (Aslin et al., 1996). The frequency of words in isolation in IDS significantly predicted children's later vocabulary knowledge (Brent & Siskind, 2001). The correlations between frequency of isolated words in IDS and children's real-world learning effects hold in Swingley and Humphrey's (2018) re-analysis of the Brent-Siskind Corpus when possible confounding factors, for example, duration of words and syntactic class, are considered. Second, placing words at utterance-final position may facilitate children's word processing compared to utterance-medial position. Fernald et al. (2001) compared English-learning children's recognition of familiar words when the words were in utterance-medial or utterance-final position, showing that 15- to 19-month-old children's word recognition performance was significantly better when they heard words produced in utterance-final position. As the authors pointed out, such facilitative effects of utterance-final position could be due to a masking effect, as utterance-medial words are masked by the rest of the utterance. Also, utterance-final words are usually longer than utterance-medial words due to final lengthening, which might be beneficial to word recognition. If mothers manipulate word position in IDS in a way that may support word learning, they would prefer placing an unfamiliar word in isolation or utterance-final position in IDS compared to ADS to magnify such beneficial effects.

The question remains whether mothers indeed prefer placing unfamiliar words in isolation or utterance-final position in IDS across languages. Fernald and Morikawa (1993) found that both American English and Japanese IDS have more words produced in isolation compared to ADS. Brent and Siskind (2001) found that 9% of utterances in American English IDS are isolated words. In American English IDS, contextually new words (focused words) are often placed in utterance-final position (Fernald & Mazzie, 1991). Similarly, Woodward and Aslin (1990) asked English-speaking mothers to teach their children unfamiliar words and found that not all mothers preferred putting unfamiliar words in isolation, but mothers placed unfamiliar words in utterance-final position 89% of the time. In a follow-up study on word order in Turkish IDS, Aslin et al. (1996) found that Turkish mothers placed unfamiliar words equally frequently at utterance-medial and utterance-final position, contrary to the findings on English IDS. It should be noted, however, that placing nouns utterance-finally is ungrammatical in Turkish, so even the lower proportion of utterance-final nouns in Turkish IDS might be viewed as surprisingly high. Perhaps Turkish mothers sacrificed grammaticality to place more words utterance finally.

The current study

In light of the possible variations in IDS speaking rate 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. In addition, we explored the position of unfamiliar words in IDS in these two languages. We had four 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; and (4) Do mothers prefer placing unfamiliar words in isolation or utterance-final position in IDS? We predict that mothers would put unfamiliar words significantly more frequently in isolation or utterance final position in IDS compared to ADS.

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). 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 et al., 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 et al., 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. Specifically, we predict a main effect of Familiarity as well as an interaction of Condition (ADS/IDS) and Familiarity (Familiar/Unfamiliar).

To address the four research questions, we conducted two experiments on Dutch (Experiment 1) and Mandarin Chinese (Experiment 2) using similar materials and speech elicitation methods. We used a picture book to elicit semi-spontaneous ADS and IDS following Fernald and Mazzie (1991). Doing so allows us to trigger target words as part of a narrative in both ADS and IDS and maintain a comparable speech context between the two speech registers. 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]

Methods

Experiment 1: Dutch

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

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

Netherlands. All children were typically developing and none of them had any hearing problems or known developmental delays.

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 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 Han (2019, p. 187) for the picture book). An additional six pages were used as fillers to make the story coherent throughout the book.

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 et al., 2009) versions of MacArthur-Bates Communicative Development Inventories (CDI) (Fenson et al., 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. The checklist resembles N-CDI. For each target word, we asked the participant mother to mark whether their child had "understood" (*begrijpen*) or "understood and said" (*begrijpen en zeggen*) it before the experiment. The score for "understood" was coded as Familiarity (Familiar/Unfamiliar) and used in data analyses.

Procedure

Each participant came to the lab twice, once at 18 months and once 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

Table 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"

[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 et al. (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.

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.

Experiment 2: Mandarin Chinese

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;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.[4] 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.[5] All children were typically developing and none of them had any hearing problems or known developmental delays.

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 Han, 2019, p. 192) for example, pages of the picture book). The Mandarin Chinese participants were tested in a quiet room in a kindergarten or in their home. The procedure was identical to Experiment 1; however, the experimenter was a native Mandarin Chinese speaker (female).

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 200 ms 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. Two additional coders annotated the word boundaries independently for 10% of the data to determine the intercoder reliability of word duration. The median and 75th percentile of word duration differences between coders were 0.02 (s) and 0.05 (s) for Dutch. The median and 75th percentile of word duration differences between coders were 0.01 (s) and 0.02 (s) for Mandarin Chinese. There were no significant differences in word duration between coders for Dutch ($t(279) = -1.43$, $p > .1$) or Chinese ($t(489) = -0.22$, $p > .1$). The annotators also transcribed and manually counted the numbers of phonological syllables for the target utterances. Two native speakers counted the numbers of syllables for 10% of the recordings; the intercoder reliability was 1.0 for Chinese and 0.93 for Dutch. In total, 1521 Dutch utterances and 1375 Chinese utterances were included for

[4]Undergraduate degree and above.

[5]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.

Table 2. Measurements on target utterances and formulas.

Measurements on target words and utterances	Formulas
Target word articulation rate (syllables/s)	$\frac{\text{Number of syllables in target word}}{\text{Target word duration}}$
Target utterance articulation rate (syllables/s)	$\frac{\text{Number of syllables in target utterances}}{\text{Total utterance duration} - \text{duration of silent pauses}}$

Table 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)

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 articulation rate in the target utterances (see Table 2 for the formulas of each measurement).

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 et al., 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 et al., 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[6]) (Bates, Kliegl et al., 2015). Then, we used the "step" function in the lmerTest package (Kuznetsova et al., 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 3. 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.

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

Results

Experiment 1: Dutch

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 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 ($p < .001$) and a significant main effect of Familiarity ($p < .001$) (Table 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 ($p < .001$), 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 that the unfamiliar words were slower than familiar words. However, the degree of slowing down in IDS did not differ between familiar and unfamiliar words.

If the unfamiliar words were not particularly slowed down in IDS compared with familiar words, how about the utterances containing unfamiliar words (see Figure 1 for box plots of Dutch utterance articulation rate)? The final model for Dutch utterance articulation rate (Table 4b) showed that there was a significant main effect of Condition ($p = .008$) as well as a significant interaction of Condition and Familiarity ($p = .022$). These results indicated that at the utterance level, the articulation rate was slowed down in IDS, and indeed the degree of slowing down was larger for utterances containing unfamiliar words in comparison with utterances containing familiar words. As the factor "Age" was

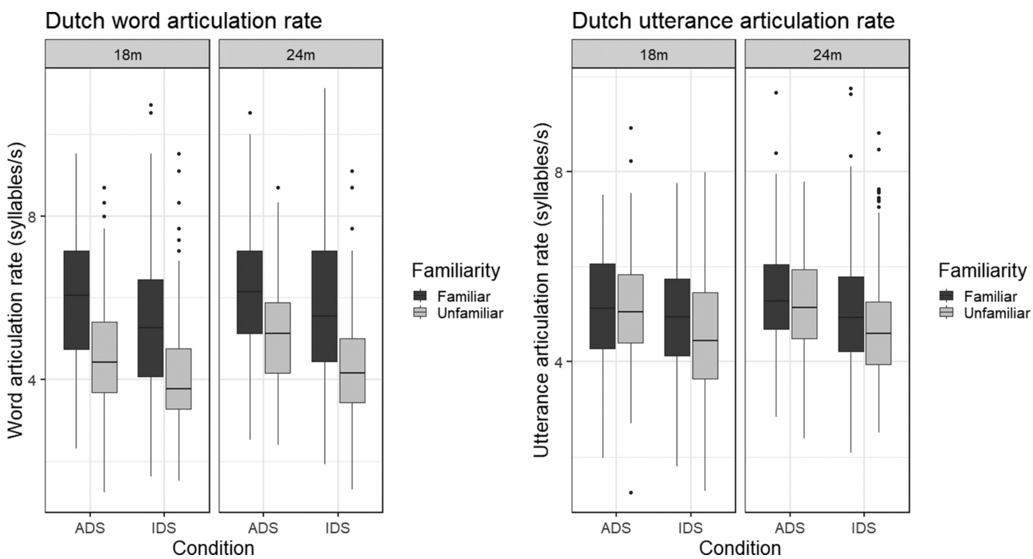

Figure 1. Box plots of word articulation rate (left panel) and utterance articulation rate (right panel) for ADS and IDS in Dutch.[8]

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.

We then explored whether mothers put unfamiliar words significantly more frequently at utterance-final position than familiar words in IDS. As there were no significant interactions between Age and

Table 4. The final models for Dutch target word articulation rate (Table 4a) and utterance articulation rate (Table 4b).

Table 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 (24 m)	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 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		

For Table 4a, Intercept represents ADS, 18 months, and Familiar. For Table 4b, Intercept represents ADS and Familiar. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Condition in the previous analyses, we collapsed the two ages for this analysis. Figure 2 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 net 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.[7] 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 ($\beta = 1.849$, $SE = 0.505$, $p < .001$), 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 ($\beta = -0.564$, $SE = 0.234$, $p = .016$), 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 place 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.

The results on articulation rate demonstrated that mothers specifically slowed down the utterances with unfamiliar words in IDS. As Dutch mothers placed target words (including familiar and unfamiliar words) more frequently in isolation in IDS than in ADS, it is possible that words in isolation, which tend to be slower than long utterances, contributed to the results. Thus, we performed additional analyses by including only long utterances (utterance syllable number > 6, utterance

[8] For all the box plots in this paper: 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.

[7] 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
```

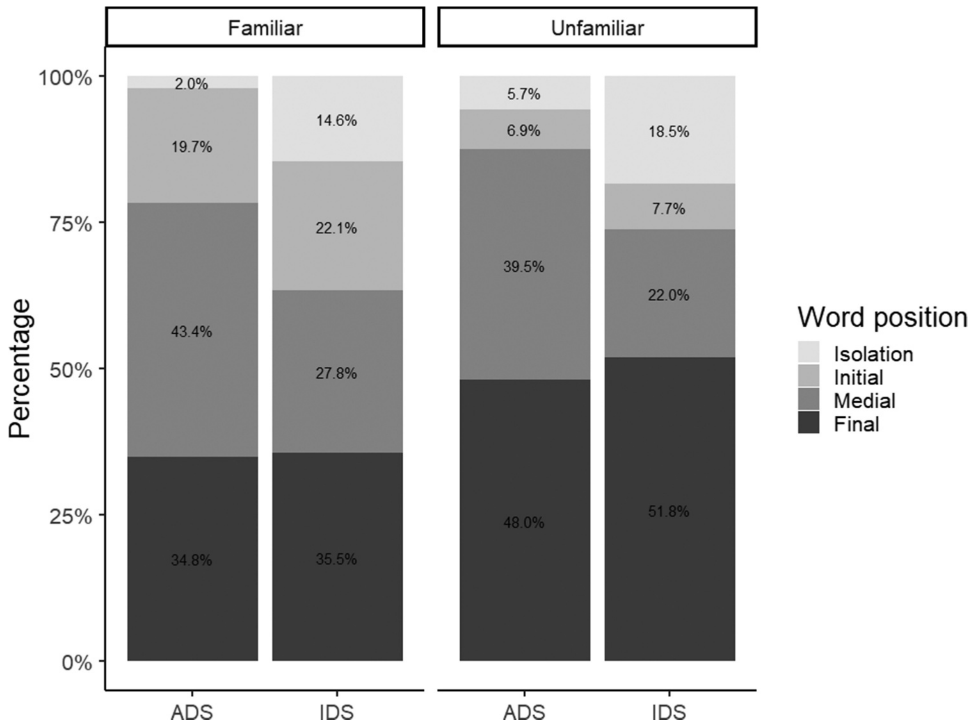

Figure 2. Word position of the target words in Dutch.

N = 970). The results showed that there was no significant effect of Condition (p = .99) or Familiarity (p = .28), but there was a significant interaction of Condition and Familiarity (β = −0.312, SE = 0.13, p = .017), suggesting that the articulation rate of long utterances, just like the effect reported for all utterances, was specifically slower for unfamiliar words in IDS. As such, it is unlikely that the results on articulation rate could be attributed to words in isolation.

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 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. As for word position, our results suggest that both familiar and unfamiliar words occurred more frequently in isolation in IDS compared to ADS but less frequently in utterance-medial position.

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 3. 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 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 = .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

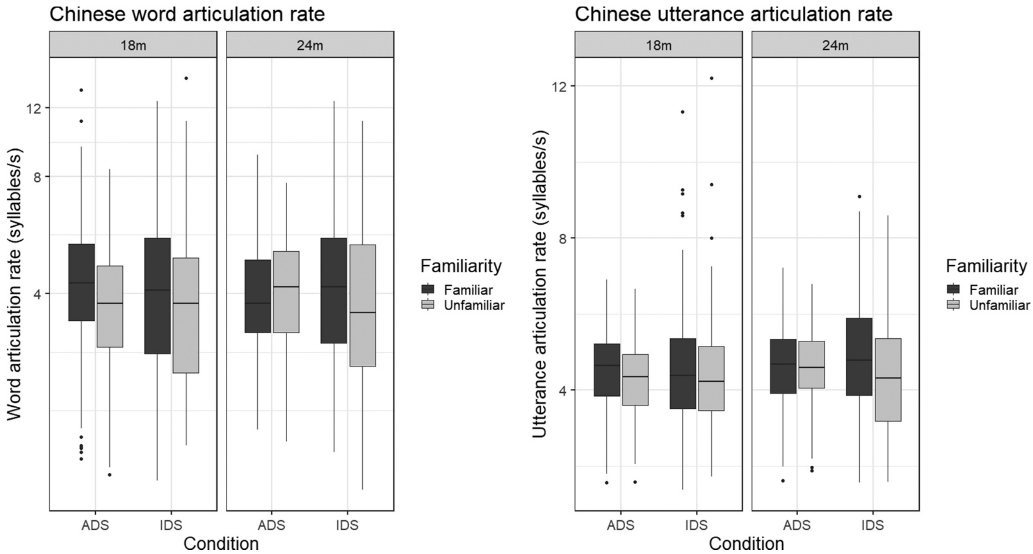

Figure 3. Box plots of word articulation rate (left panel, y-axis is log-transformed) and utterance articulation rate (right panel) for ADS and IDS in Mandarin Chinese.

(Table 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 3. 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 = .036), indicating that utterances with unfamiliar words were slower than utterances with familiar words in both ADS and IDS.

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 4 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 only a significant effect of Condition on isolation position (β = 1.398, SE = 0.399, p < .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 target words more often in isolation but do not put target words more often in utterance-final position in Mandarin Chinese IDS, and this trend holds for both familiar and unfamiliar words.

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.

Discussion

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

Table 5. The final models for Mandarin Chinese target word articulation rate (log- transformed) (Table 5a) and utterance articulation rate (Table 5b).

Table 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 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		

Intercept represents Familiar in both Table 5a and Table 5b * p < 0.05; ** p < 0.01; *** p < 0.001.

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. 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 & McDermott, 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 their 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 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.

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? Han et al. (2020) examined the pitch properties of the same speech corpora and found that mothers did modify pitch in Mandarin Chinese IDS. 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 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

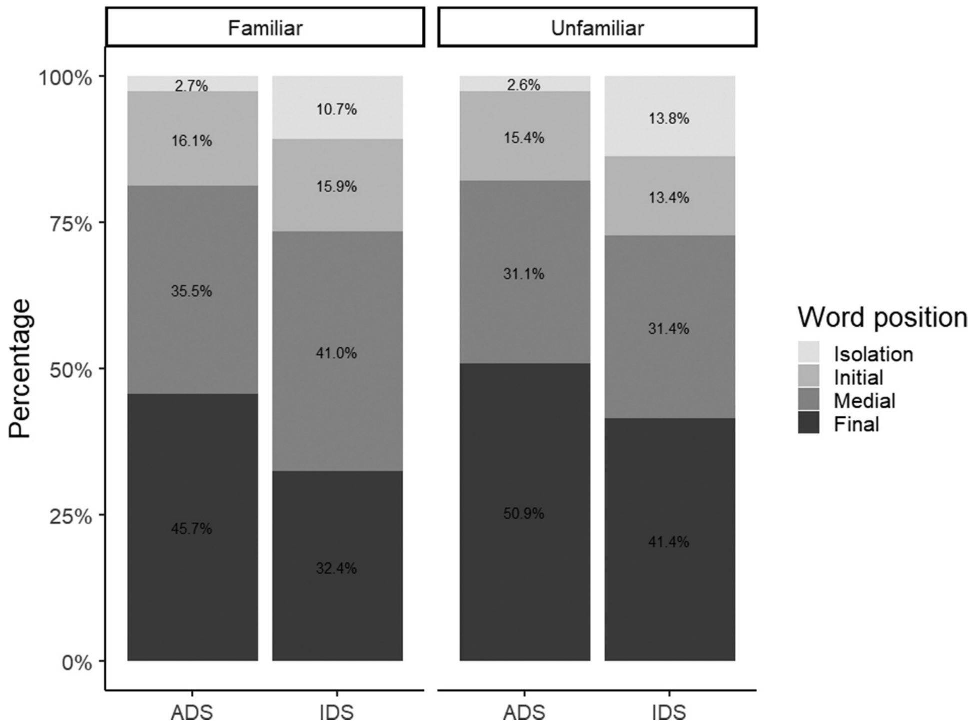

Figure 4. Word position of the target words in Mandarin Chinese.

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 et al., 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 fewer 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. 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 & McDermott, 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 languages), 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. In contrast, Mandarin Chinese uses pitch-related cues such as pitch height and pitch range instead to mark unfamiliar words (Han et al., 2020).

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, most 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 et al., 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 et al., 2020; 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 position in IDS. Fernald and Mazzie (1991) reported that English IDS tends to put focused words in utterance-final position. Even though this finding has often been referred to as a major characteristic of IDS word position, 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. Also, mothers did not manipulate word position to highlight unfamiliar words in IDS. The finding that both Dutch and Mandarin Chinese mothers produced more target words in isolation IDS than ADS are consistent with previous studies on English and Japanese (Fernald & Morikawa, 1993), which are helpful for children's word segmentation (Aslin et al., 1996). Fernald et al. (2001) indicate that there are facilitative effects of utterance-final position on English-learning children's word recognition. However, whether such effects may be attributed to utterance final-lengthening, which is a speech phenomenon that is found in many languages such as English, Dutch, and Japanese (Cambier-Langeveld, 1997; Downing, 1970; Martin et al., 2016), or to the fact that utterance-final words are not masked, or it is driven by the language-specific word position preference in English IDS, requires further investigation.

One question arises as to why Mandarin Chinese IDS had fewer words in utterance-final position than ADS, while Dutch IDS had fewer words in utterance-medial position compared to ADS. A possible explanation may be attributed to the canonical word order difference between the two languages. The target words (nouns) can each be used as either a subject or an object in a sentence. Since Mandarin Chinese is an SVO language (Sun & Givón, 1985) and Dutch is and SOV language (Koster, 1975), it's possible that mothers in both language groups used target words as objects less often in IDS than in ADS, leading to differences in word order preferences between the two languages. Future studies may investigate the syntactic structure and sentence type of the sentences embedding the target words as well as the influence of pragmatics on the choice of word positions.

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 be confounded 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, 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 would 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 lower in frequency) 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 et al. (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.

In this study, we used a picture book to elicit semi-spontaneous ADS and IDS following Fernald and Mazzie (1991). This speech elicitation method ensures that the speech context is comparable between ADS and IDS. However, this specific task may not be generalizable to other speech contexts in naturalistic settings. Thus, we compared our results for speaking rate with natural Dutch and Mandarin corpora. In the IFA Dutch Spoken Language Corpus, the mean articulation rates (excluding pauses) were 5.5 syllables/s for "informal story telling face-to-face to an 'interviewer'" and 5.2 syllables/s for "retelling a previously read narrative story without sight contact" (Van Son et al., 2001). Our results showed that the utterance articulation rates of Dutch ADS are all around 5.2 syllables/s, which are similar to those in the IFA corpus. To the best of our knowledge, no study has systematically investigated the speaking rate of Mandarin Chinese (semi-) spontaneous speech. In a cross-linguistic study, Pellegrino et al. (2011) found that the articulation rate of Mandarin Chinese read speech (with scripts) was around 5.18 syllables/second. Moreover, their results suggest that Mandarin Chinese is the slowest among all the languages under investigation, which could explain the difference in articulation rate between Dutch and Chinese ADS in our data. In another Mandarin Chinese read speech corpus (Chiang et al., 2009), the normal articulation rate of Mandarin Chinese is 5.05 syllables/s and the median articulation rate is 4.15 syllables/s. Our results showed that the Mandarin Chinese utterance articulation rates are around 4.5 syllables/second. As read speech is often faster than (semi-) spontaneous speech (see Van Son et al., 2001), overall, our results of Chinese ADS articulation rate are comparable to other speech corpora.

Conclusion

Despite robust evidence supporting the universality of IDS, our results reveal language-specificity in speaking rate and word position of IDS in word-learning contexts between Dutch and Mandarin Chinese – two typologically distinct languages. Specifically, Dutch IDS was generally slower than 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 always be slower than ADS. In both languages, mothers put target words in isolation more frequently in IDS in both languages, but they did not prefer putting target words in utterance-final position, inconsistent with previous findings on English IDS. 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.

Acknowledgments

We are grateful to Taohualing Kindergarten (Mrs. Aihua Zou and Mrs. Yuhong Li), Gezhouba Dongshan Kindergarten (Mrs. Huan He) and Gezhouba Early Education Center (Mrs. Hong Xie) in Yichang, China for their kind support and coordination in recruiting participants. We thank Run Chai, Lisanne Geurts, Karlijn Kouwer, Anika van der Klis, Judith van der Linden, and Anqi Yang for their help with data collection and annotation. We also acknowledge the members of Utrecht Babylab, Hans Rutger Bosker, Jeroen Breteler, and Yan Gu for their input. We thank all the families who participated in this study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Shanghai Planning Project of Philosophy and Social Sciences (2020EYY001); Universiteit Utrecht.

ORCID

Nivja H. de Jong http://orcid.org/0000-0002-3680-3820
René Kager http://orcid.org/0000-0002-5811-839X

References

Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 117–134). LEA.

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

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*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). *Parsimonious mixed models.* Available from arXiv:1506.04967 (stat.ME) 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. https://doi.org/10.18637/jss.v067.i01

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. https://doi.org/10.1016/j.jml.2008.06.003

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. https://doi.org/10.1073/pnas.1113380109

Bernstein-Ratner, N. (1985). Dissociations between vowel durations and formant frequency characteristics. *Journal of Speech, Language, and Hearing Research*, *28*(2), 255–264. https://doi.org/10.1044/jshr.2802.255

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

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

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33–B44. https://doi.org/10.1016/S0010-0277(01)00122-6

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

Cambier-Langeveld, T. (1997). The domain of final lengthening in the production of Dutch. *Linguistics in the Netherlands*, *14*(1), 13–24. https://doi.org/10.1075/avt.14.04cam

Chiang, C. Y., Tang, C. C., Yu, H. M., Wang, Y. R., & Chen, S. H. (2009, August). An investigation on the mandarin prosody of a parallel multi-speaking rate speech corpus. In *2009 oriental COCOSDA international conference on speech database and assessments* (pp. 148–153). IEEE.

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* (2017), 27–34. https://doi.org/10.1016/j.cognition.2016.12.002

Downing, B. T. (1970). *Syntactic structure and phonological phrasing in English* [Unpublished doctoral dissertation]. University of Texas at Austin.

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*. Brookes.

Fernald, A. (2000). Speech to infants as hyperspeech: Knowledge-driven processes in early word recognition. *Phonetica*, *57*(2–4), 242–254. https://doi.org/10.1159/000028477

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., McRoberts, G. W., & Swingley, D. (2001). Infants' developing competence in recognizing and understanding words in fluent speech. In J. Weissenborn & B. Hoehle (Eds.), *Approaches to Bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition* (Vol. 1, pp. 97–123). Benjamins.

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.

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*(1), 171–183. https://doi.org/10.1017/S0305000900013167

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 (Eds.), *Papers in laboratory phonology* (pp. 515–546). Mouton de Gruyter.

Han, M. (2019). *The role of prosodic input in word learning: A cross-linguistic investigation of Dutch and Mandarin Chinese infant-directed speech* [Doctoral dissertation, Utrecht University]. LOT Publications.

Han, M., De Jong, N. H., & Kager, R. (2020). Pitch properties of infant-directed speech specific to word-learning contexts: A cross-linguistic investigation of Mandarin Chinese and Dutch. *Journal of Child Language*, *47*(1), 85–111. https://doi.org/10.1017/S0305000919000813. https://doi.org/10.1017/S0305000919000813

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., 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

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)

Koster, J. (1975). Dutch as an SOV language. *Linguistic Analysis*, *1*(2), 111–136.Retrieved from: https://www.dbnl.org/tekst/kost007dutc01_01/kost007dutc01_01.pdf
Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*(5326), 684–686. https://doi.org/10.1126/science.277.5326.684
Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4
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. https://doi.org/10.18637/jss.v082.i13
Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2017). The temporal modulation structure of infant-directed speech. *Open Mind*, *1*(2), 78–90. https://doi.org/10.1162/OPMI_a_00008
Li, D. C., & Lee, S. (2006). Bilingualism in East Asia. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 742–779). Blackwell Publishing Ltd. https://doi.org/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. https://doi.org/10.1177/002383096300600306
Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Kluwer.
Martin, A., Igarashi, Y., Jincho, N., & Mazuka, R. (2016). Utterances in infant-directed speech are shorter, not slower. *Cognition*, *156*(2016), 52–59. https://doi.org/10.1016/j.cognition.2016.07.015
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. https://doi.org/10.1016/j.cognition.2013.07.015
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 (Eds.), *Proceedings of the 4th conference on speech prosody* (pp. 423–426). 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. https://doi.org/10.1121/1.4944634
Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language*, *87*(3), 539–558. https://doi.org/10.1353/lan.2011.0057
Quené, H. (2007). On the just noticeable difference for tempo in speech. *Journal of Phonetics*, *35*(3), 353–362. https://doi.org/10.1016/j.wocn.2006.09.001
R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
Ramus, F., Dupoux, E., & Mehler, J. (2003). *The psychological reality of rhythm classes: Perceptual studies* [Paper presentation]. 15th International Congress of Phonetic Sciences, Barcelona, Spain.
Raneri, D., Von Holzen, K., Newman, R., & Bernstein Ratner, N. B. (2020). Change in maternal speech rate to preverbal infants over the first two years of life. *Journal of Child Language*, *47*(6), 1263–1275. https://doi.org/10.1017/S030500091900093X
Seifart, F., Strunk, J., Danielsen, S., Hartmann, I., Pakendorf, B., Wichmann, S., Witzlack-Makarevich, A., de Jong, N. H., & Bickel, B. (2018). Nouns slow down speech across structurally and culturally diverse languages. *Proceedings of the National Academy of Sciences*, *115*(22), 5720–5725. https://doi.org/10.1073/pnas.1800708115
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). Red Hook, NY: Curran Associates, Inc.
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. https://doi.org/10.1121/1.417955
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
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
Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspec ts of infant-directed speech and infant outcomes. *Developmental Review, 44,* 1–18. https://doi.org/10.1016/j.dr.2016.12.001
Sun, C. F., & Givón, T. (1985). On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications. *Language*, *61*(2), 329–351. https://doi.org/10.2307/414148
Swingley, D., & Humphrey, C. (2018). Quantitative linguistic predictors of infants' learning of specific English words. *Child Development*, *89*(4), 1247–1267. https://doi.org/10.1111/cdev.12731
Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2018). Routine language: Speech directed to infants during home activities. *Child Development*, *90*(6), 2135–2152. 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.Retrieved from https://www.phon.ucl.ac.uk/home/shl9/tang/tangjm.htm

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.

Van Son, R. J. J. H., Binnenpoorte, D., van den Heuvel, H., & Pols, L. C. W. (2001). The IFA corpus: A phonemically segmented Dutch "open source" speech database. *EUROSPEECH*, *2001*, 2051–2054. https://repository.ubn.ru.nl/bitstream/handle/2066/76439/76439.pdf

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.

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 (Eds.), *Dimensions of phonological stress* (pp. 311–324). Cambridge University Press.

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. https://doi.org/10.1121/1.4979704

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. https://doi.org/10.1016/S0163-6383(84)80022-3

Woodward, J. Z., & Aslin, R. N. (1990, April). *Segmentation cues in maternal speech to infants* [Presentation]. International Conference on Infant Studies, Montreal.

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. https://doi.org/10.1207/s15327647jcd0602_2

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. https://doi.org/10.1207/s15327078in1101_2

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