Prosodic input and children’s word learning in infant- and adult-directed speech

Han, M.; Jong, N.H. de; Kager, R.

Infant Behavior and Development 68 (2022) 101728

Contents lists available at ScienceDirect

Infant Behavior and Development

journal homepage: www.elsevier.com/locate/inbede

Prosodic input and children's word learning in infant- and adult-directed speech

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

[a] Department of Chinese Language and Literature, East China Normal University, 500 Dongchuan Road, 200241 Shanghai, China
[b] Utrecht Institute of Linguistics (OTS), Utrecht University, Trans 10, 3512 JK Utrecht, the Netherlands
[c] Language, Cognition, and Evolution Lab, East China Normal University, 500 Dongchuan Road, 200241 Shanghai, China
[d] Leiden University Center for Linguistics (LUCL), Leiden University, Van Wijkplaats 4, 2311 BX Leiden, the Netherlands
[e] Leiden University Graduate School of Teaching (ICLON), Leiden University, Kolffpad 1, 2333 BN Leiden, the Netherlands

ARTICLE INFO

Keywords:
Infant-directed speech
Maternal input, prosody
Word-learning

ABSTRACT

This study examines (1) whether infant-directed speech (IDS) facilitates children's word learning compared to adult-directed speech (ADS); and (2) the link between the prosody of IDS in word-learning contexts and children's word learning from ADS and IDS. Twenty-four Dutch mother-child dyads participated when children were 18 and 24 months old. We collect mothers' ADS and IDS at both ages and test children's word learning from ADS and IDS at 24 months. We find that Dutch 24-month-old children could reliably learn novel words from both ADS and IDS, and IDS had a facilitative effect. In addition, children's word learning from IDS (but not ADS) is predicted by IDS pitch range when mothers introduce unfamiliar words to children at 18 months. Our findings contribute to an understanding of the role of IDS prosody in language development, highlighting both individual differences and contextual differences in IDS prosody.

1. Introduction

When addressing young children, mothers modify their speech prosody by heightening their pitch level, expanding their pitch range, and slowing down their speaking rate (see reviews in Cristia, 2013; Soderstrom, 2007). This speaking register, known as infant-directed speech (IDS), has been widely observed in many languages around the world. Even though IDS has long been proposed to facilitate children's language acquisition compared to adult-directed speech (ADS) (see a review in Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015), the role of IDS prosody in children's word learning has not been fully addressed in the literature: First, despite some evidence that children learn novel words better from prototypical American English IDS compared to ADS (Graf Estes & Hurley, 2013; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011), it is unclear whether this holds across languages. Second, it remains a question whether children's word learning performance in IDS and ADS is related to individual differences in the prosodic input they are exposed to.

The overarching aim of the current study is to further examine the role of IDS prosody in children's word learning. We first test

* Corresponding author at: Department of Chinese Language and Literature, East China Normal University, 500 Dongchuan Road, 200241, Shanghai, China.
** Correspondence to: Leiden University, Leiden University Centre for Linguistics (LUCL) & Leiden University Graduate School of Teaching (ICLON), the Netherlands.
E-mail addresses: mrhan@zhwx.ecnu.edu.cn (M. Han), n.h.de.jong@hum.leidenuniv.nl (N.H. De Jong).

https://doi.org/10.1016/j.infbeh.2022.101728
Received 20 February 2022; Received in revised form 27 May 2022; Accepted 27 May 2022
Available online 14 June 2022

whether prototypical IDS facilitates 24-month-old Dutch children's novel word learning compared to ADS word learning in a laboratory setting. We also explore whether there is a link between individual mothers' prosodic input and children's word learning from ADS and IDS. Specifically, we focus on IDS prosody in word-learning contexts in which mothers introduce unfamiliar words to their children.

1.1. Effects of prototypical IDS on novel word learning

A number of studies have shown that prototypical IDS facilitates children's word processing compared to ADS. For example, English- and German-learning infants could only segment words from continuous speech when hearing prototypical IDS but not ADS (Mani & Pätzold, 2016; Thiessen, Hill, & Saffran, 2005). English-learning infants recognized words they were familiarized with in prototypical IDS, but not words that were introduced in ADS (Singh, Nestor, Parikh, & Yull, 2009). Regarding novel word learning, two studies have shown that prototypical IDS facilitates novel word learning in American-English-learning children. Ma et al. (2011) tested whether 21- and 27-month-old American-English-learning children learned novel words better from IDS compared to ADS using an Intermodal Preferential Looking Paradigm (IPLP) (Hirsh-Pasek & Golinkoff, 1996). For the 21-month-old group, they adopted a between-subject design, and children were randomly assigned to ADS or IDS conditions. For the 27-month-old group, they only tested children's word learning from ADS. The findings suggest that 21-month-old children could reliably learn novel word-to-object mappings when the words were taught in IDS but not in ADS, but 27-month-old children could reliably learn novel words from ADS. As they only tested the ADS condition at 27 months, it is unclear whether IDS still had a facilitative effect in word learning even when children could learn words from ADS at that age. In addition to the group differences, they found individual differences in children's word learning: Twenty-one-month-old children who had a relatively larger vocabulary size performed better in the word-learning task compared to children with smaller vocabularies. Also, only 21-month-old children with a larger vocabulary could learn novel words in ADS. In another study, Graf Estes and Hurley (2013), using a Switch task (Werker, Cohen, Lloyd, Casasola & Stager, 1998), found that 17-month-old children could learn word-object associations in the IDS condition but not in the ADS condition.

Taken together, findings from these two studies suggest that American English IDS has a facilitative effect on word learning at least until children reach 21 months of age. However, it should be noted that the degree of prosodic exaggeration in IDS varies among languages: American English IDS is more exaggerated compared to British English, French, Italian, German, and Japanese IDS (Fernald et al., 1989), and Thai IDS is less exaggerated than Australian English IDS (Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002). Such differences in the degree of prosodic exaggeration in IDS may further influence children's word processing. 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 were particularly exaggerated in prosody (Floccia et al., 2016). Thus, even though evidence shows that American English IDS facilitates children's word learning, whether this conclusion can be generalized to other languages in which IDS was less exaggerated than American English remains a question.

1.2. Individual differences in children's word learning in relation to input quality

Studies that compared children's word processing in ADS and IDS often examined children's learning performance in a laboratory setting independent from their maternal input, focusing on group differences instead of individual differences in input quality. The question remains whether children's word learning from IDS and ADS is related to individual differences in the prosodic input they are exposed to.

Even though exaggerated prosody is the most prominent feature of IDS, and prototypical IDS facilitates children's word learning compared to ADS (at least for American English), studies on the link between IDS prosody and children's online word processing often yield null results. For example, Suttora et al. (2017) explored the relationship between the prosody of IDS (including mean F0, maximum F0, minimum F0, and speech rate) and Italian-learning children's word recognition at 15 months. Their results reveal no significant concurrent correlations between IDS prosody and children's word recognition accuracy. Similarly, Song, Demuth, and Morgan (2018) did not find any correlations between the prosody of American English IDS (including mean pitch and pitch range) at 17 months and children's vocabulary size at 19 or 25 months. In a recent study, Outters, Schreiner, Behne and Mani (2020) found a link between the prosody of IDS and children's listening preference for IDS compared to ADS. In particular, mothers' higher mean pitch in IDS (relative to ADS) is related to a greater listening preference for IDS at 6 months but not at 13 months. Even though children's listening preferences do not necessarily lead to better learning, it is possible that the prosodic exaggeration in IDS relative to ADS can also be a predictor of children's word learning from IDS.

1.3. Contextual differences in prosodic input

In the correlational studies reviewed above, the predictors were mostly global measures of IDS prosody during mother-child interactions. Some evidence suggests that the prosody of IDS shows contextual differences. For example, Fernald and Mazzie (1991) found that American-English-speaking mothers used prosody to highlight contextually new words compared to contextually given words. Han, de Jong, and Kager (2020), (2021) found that Dutch and Chinese mothers modified their speech prosody when introducing familiar vs. unfamiliar words to children. Dutch mothers had a lower pitch and a slower articulation rate when introducing unfamiliar words compared to familiar words, while Mandarin-Chinese-speaking mothers heightened pitch and expanded pitch range. Word-learning contexts provide the most direct input for children to obtain information for word learning. In a recent study, Van Rooijen, Bekkers, and Junge (2019) tested 24-month-old Dutch children's word learning from their mother's speech and an experimenter's speech. In this experiment, the content of the speech stimuli was the same for the mother and the experimenter, and both speakers were instructed to introduce the novel word labels using natural IDS speech. Still, children learned novel words faster from their mothers' speech compared to the experimenter's speech. This points to an intriguing possibility that the prosody of an individual mother's IDS in word-learning contexts could be correlated with children's word learning performance.

1.4. The current study

To summarize, whether the facilitative effects of IDS on children's novel word learning is language-universal still requires evidence from languages other than American English. Also, it remains a question whether the facilitative effects of IDS are robust even when children can reliably learn novel words from ADS. In addition, no study has investigated the link between prosodic input in word-learning contexts and children's word learning from ADS and IDS. The overarching aim of the current study is to further examine the link between the prosody of individual mothers' IDS and children's word learning.

In the present study, we addressed two research questions. First, does prototypical IDS facilitate 24-month-old Dutch children's novel word learning compared to ADS? To answer this question, we conceptually replicated the American English word learning experiment by Ma et al. (2011) and tested Dutch children's novel word learning from ADS and IDS using a within-subject design. We predicted that Dutch IDS has a facilitative effect on word learning compared to ADS.

Second, is there a link between individual mothers' IDS prosody and children's word learning from ADS and IDS? As word-learning contexts, in which mothers introduce novel words to children, are immediately relevant to children's word learning, we focused on mothers' IDS prosody in such contexts and examined whether it is related to children's novel word learning from ADS and IDS. There are two possible hypotheses: (1) If children's word learning is related to their experience with IDS, their primary prosodic input, a mother's IDS prosody would be correlated with their child's word learning performance in the IDS condition only; Alternatively, (2) If IDS facilitates children's word learning ability in general, a mother's IDS prosody would be similarly correlated with word learning in both ADS and IDS.

Individual differences in IDS prosody can be evaluated in two ways. In most studies, the raw prosodic values of each mother's IDS are used as predictors. These values indicate the IDS prosody children are exposed to. Kalashnikova and Burnham (2018) calculated a hyper-score for each mother, dividing each prosodic value in IDS by their corresponding ADS values. These hyper-scores indicate the extent to which each mother's IDS prosody differs from their ADS. We adopt two sets of prosodic predictors for each mother: raw prosodic values of IDS and prosodic hyper-scores.

To answer the above two questions, we tested the same children on their word learning from ADS and IDS at 24 months and collected their mothers' ADS and IDS at 18 and 24 months. We targeted these two ages as children's vocabulary increases rapidly and children's fast mapping improves greatly during this period (Bion, Borovsky, & Fernald, 2013; Goldfield & Reznick, 1990). By focusing on this period, we may examine how prosodic input contributes to this process.

2. Method

2.1. Participants

This study is part of a larger study on the role of prosodic input in word learning (Han, 2019). Participants include twenty-four Dutch mother-child dyads who visited the lab when the children were aged 18 and 24 months (girls $N = 11$; mean age of mothers $= 35$ years, mothers' age range $= 31$–44 years; Mean age of testing at 18 m: $= 18;12$ (age range $= 18;00$–18;29); Mean age of testing at 24 m $= 24;20$, age range $= 24;00$–27;00). The participants were recruited from the Utrecht Baby Lab database and are all Dutch native speakers living in the Utrecht area in the Netherlands. All children were typically-developing and without reported language impairments or hearing problems. Written informed consent was obtained for all participants.

2.2. General procedure

The same mother-child dyads visited the lab three times, once at 18 months and twice after the children reached 24 months. During the first visit at 18 months, we collected mothers' speech addressed to their child (IDS) and an adult (ADS). Participants visited the Utrecht Baby Lab twice on two different days after they reached 24 months (24;00). We collected participant mothers' ADS and IDS addressed to 24-month-old children on the first visit at 24 months. Participants were also tested on word learning from ADS or IDS separately during the two visits (order counterbalanced). The average duration between the two visits was 7.6 days (range $= 2$–35 days).

2.3. Speech data collection

2.3.1. Materials

The speech data collection methods were identical to those reported in Han et al., (2020), (2021). We created two books for the two respective age groups and used a storybook-telling task to elicit semi-spontaneous speech from the mothers. The two books were identical in structure. Each book contained a set of target words that were unfamiliar to children (see Supplementary Table 1 for a list of the preselected unfamiliar words). 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 for the picture books).

2.3.2. Procedure

During each lab visit, mothers told the story twice. For ADS, mothers were instructed to tell the story to an experimenter (a female Dutch native speaker), and to take into account the fact that she was a college student. For IDS, the mother was instructed to tell the story to her child the way she normally would at home. After each experiment, mothers filled out a word checklist to determine whether their child had already known the target words before the experiment. As we focused on word-learning contexts in which mothers introduced unfamiliar words to children, we only included the words that mothers marked "unfamiliar" in analyses.

2.3.3. Data analysis

We measured the prosody of utterances containing unfamiliar words ($N = 636$) using a Praat script. Specifically, we measured articulation rate (syllables/s), mean F0 (in semitones (st)), and F0 range (Maximum F0 – Minimum F0, in semitones (st)[1]). We performed repeated measures ANOVAs to examine the effect of Age (18 m/24 m) and Condition (ADS/IDS) on these prosodic measures. The results show a significant main effect of Condition on articulation rate ($F(1, 17) = 17.24$, $p < 0.001$) and mean pitch ($F(1, 17) = 17.24$, $p = 0.006$), but not on pitch range ($F(1, 17) = 1.32$, $p = 0.27$). There was no significant interaction between Age and Condition (all p's > 0.1). These results suggest that in this corpus, Dutch mothers had a slower articulation rate and a higher pitch in IDS than ADS when children were 18 and 24 months old.

2.4. Word learning experiment

The word learning experimental procedure was adapted from Ma et al. (2011). We adopt a within-subject design instead of the between-subject design used in their study.

2.4.1. 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 screen's height 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 fixations. Audio stimuli were delivered through a speaker (Tangent EVO-E4) in front of the testing area.

2.4.2. 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 Supplementary Table 2).

A female Dutch native speaker recorded the audio stimuli. For the ADS condition, the speaker was asked to imagine that she was talking to an adult. For the IDS condition, she was asked to imagine that she was talking to an infant. The speaker recorded five versions for each speech register. Three Dutch native speakers who were blind to the speech registers of each recording judged the speech register (Condition: ADS/IDS) as well as the naturalness of the recordings. We selected the versions of ADS and IDS judged to be most natural for use as stimuli. We used Praat to extract three measures of prosody from the stimuli: mean F0 (Hz), F0 range (Hz), and articulation rate (Table 1). We used linear mixed-effects models to compare the prosodic differences between the audio stimuli in ADS and IDS. The results show significant main effects of Condition on mean F0 ($\beta = 87.837$, $SE = 5.57$, $t = 15.78$, $p < 0.001$), F0 range ($\beta = 165.38$, $SE = 31.66$, $t = 5.224$, $p < 0.001$) and articulation rate ($\beta = -0.17$, $SE = 0.43$, $\mathrm{t} = -3.93$, $p = 0.002$), suggesting that the audio stimuli in the IDS condition had a higher pitch, a larger pitch range, and a slower articulation rate compared to ADS, in line with previous studies on Dutch IDS (Benders, 2013; Van de Weijer, 1999).

2.4.3. Visual stimuli

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

2.4.4. Procedure

During the experiment, children sat on their mother's lap facing the monitor. The mothers 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

[1] Following Kitamura et al. (2002), F0 range (Hz) was transformed to Semitones (st) using the formula: Semitones = 12 *$\log_2$(maximum F0/minimum F0). Mean F0 (Hz) was transformed to Semitones (st) using the formula: 12LOG_2(Mean F0/50)

Table 1
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)

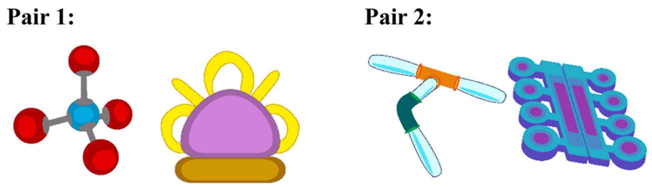

Fig. 1. Visual stimuli.

session consisted of five phases: task familiarization, salience, training, and testing phases (see Supplementary Table 2 for an overview of procedure). The order of the two conditions was counterbalanced across participants.

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

2.4.5. *Coding*

Coding was done using the UiL OTS Video Coding System (De Mooij, 2018). A coder who was blind to condition and trial coded the looking directions (left, center, right, distracted, unclear) frame-by-frame throughout each video. An independent coder coded 10% of the trials with 98.5% agreement.

2.4.6. *Data analysis*

As in Ma et al. (2011), we measured single longest looks at the target and to the distractor for each testing trial in the entire trial (6000 ms). 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 (following Van Rooijen et al., 2019). We also measured proportion of looking time (calculated as Target looking time / (Target + Distractor looking time)) and latency of shift to target from the distractor 200 ms after participants hear the target words (Delle Luche, Durrant, Poltrock, & Floccia, 2015; Fernald, Zangl, Portillo, & Marchman, 2008). For the latency measure, only the trials in which children initially looked at the distractor (74/192 trials) were included in the analysis.

3. Results

3.1. *Word learning from ADS and IDS*

Table 2 shows the means and standard deviations of the three dependent measures for Target and Distractor in each condition (ADS/IDS). We used the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in the R environment (R Core Team, 2018) for analysis. We used linear mixed-effects models to examine whether children learned novel words better in IDS than ADS. The dependent

Table 2
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 (1544)
		Distractor	1554 (1100)
	IDS	Target	2987 (1542)
		Distractor	1326 (837)
Proportion of looking time (%)	ADS	Target	65.1 (28.8)
		Distractor	57.6 (26.0)
	IDS	Target	66.6 (26.6)
		Distractor	51.2 (27.1)
Latency (ms)	ADS		836 (429)
	IDS		771 (445)

measures were single longest look (ms), proportions of looking time (%), and latency (ms). The fixed factors were 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). 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 shown in Supplementary Table 3.

To check whether children paid more attention in the IDS Condition than in the ADS condition, we examined whether children's looking time in the entire testing phase (6000 ms) was longer in either condition. The results showed no significant differences (Mean looking time in the ADS Condition = 2553 (ms), $SD = 1650$; Mean looking time in the IDS Condition = 2716 (ms), $SD = 1667$; $\beta = 182.7$, $SE = 170.2$, $p = 0.284$). Thus, there was no evidence for attentional differences in the testing phase between ADS and IDS conditions.

We then tested the effects of Condition and Target on the three dependent measures: single longest look, proportion of looking time, and latency. Fig. 2 shows box plots of children's single longest look to the target and the distractor in ADS and IDS conditions. The data were log-transformed before statistical analysis. For this measure, the final model (Supplementary Table 3a) showed a significant main effect of Target ($\beta = 0.47$, $SE = 0.10$, $t = 4.55$, $p < 0.001$) and a significant interaction between Condition and Target ($\beta = 0.37$, $SE = 0.15$, $t = 2.51$, $p = 0.013$), however, the main effect of Condition was not significant. These results show that children's single longest look at the target was longer compared to the distractor in ADS, and it was specifically longer in IDS compared to ADS. In line with Ma et al. (2011), these results suggest that IDS has a facilitative effect on children's word learning, though Dutch children could learn novel words in both ADS and IDS at 24 months.

Fig. 3 shows box plots of proportions of looking time to the target or the distractor in ADS and IDS conditions. The final model for proportions of looking time (Supplementary Table 3b) showed that there was a significant main effect of Target ($\beta = 0.28$, $SE = 0.04$, $t = 7.55$, $p < 0.001$), but there was no significant main effect of Condition, nor was there a significant interaction between Condition and Target. These results suggest that children looked significantly longer to the target compared to the distractor in both conditions. As such, children learned the novel words regardless of the speech register.

We further tested whether children were faster to look at the target words in the IDS condition compared to the ADS condition. Fig. 4 shows boxplots of latency in ADS and IDS conditions. No random factors or fixed effects remained in the final model. As such, there was no evidence to suggest that children looked at the target significantly faster in the IDS condition compared to the ADS condition.

In summary, the word learning experiment showed that Dutch 24-month-old children could already reliably learn novel words from ADS and IDS, though there was still a facilitative effect of IDS on word learning.

3.2. Correlations between IDS prosody in word-learning contexts and children's word learning from ADS and IDS

We conducted a series of multiple regression analyses using the lm() function in the R environment (R Core Team, 2018). Before building each model, we 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). For each model, we started by including all the predictors 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: "backward").

The predictors were the mean raw prosodic values and prosodic hyper-scores of each participant. The raw prosodic values were

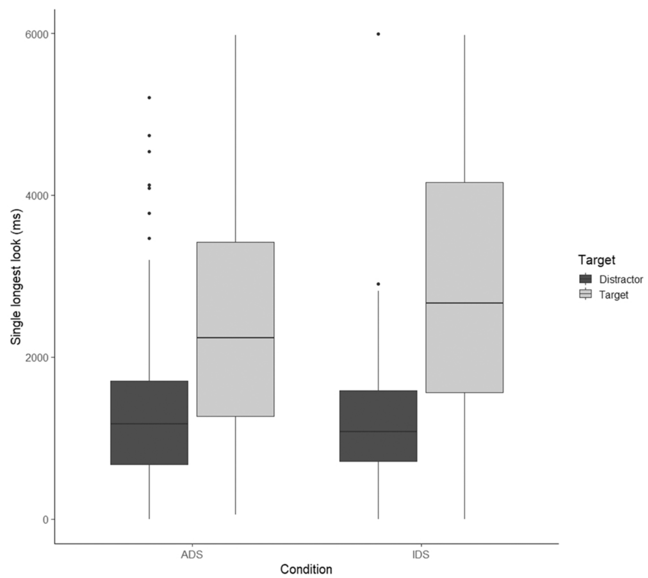

Fig. 2. Box plots of single longest look (ms) to the target and the distractor in ADS and IDS conditions.

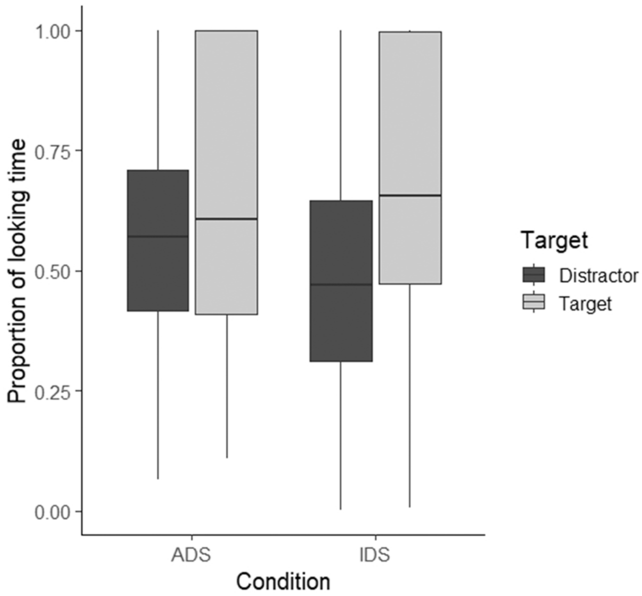

Fig. 3. Box plots of proportion of looking time to the target and the distractor in ADS and IDS conditions.

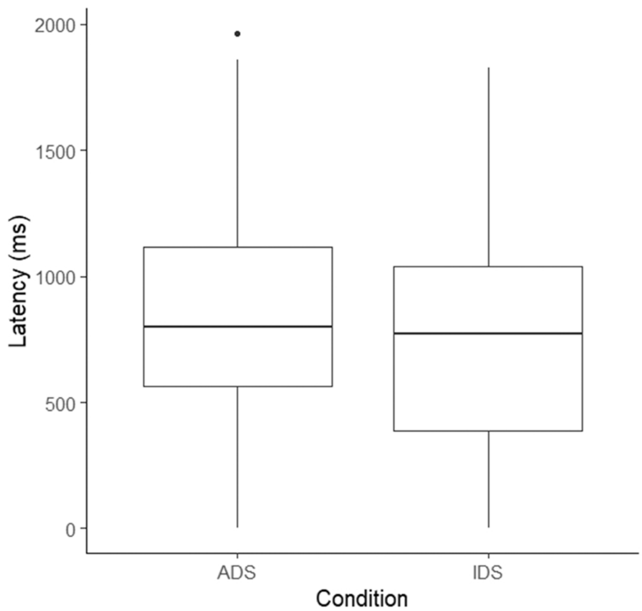

Fig. 4. Box plots of latency in ADS and IDS conditions.

calculated by averaging the prosodic measurements for each mother, indicating IDS prosody in word learning contexts. The hyper-scores were calculated by dividing raw IDS prosodic values by ADS values when mothers introduce unfamiliar words, indicating the prosodic exaggeration of IDS compared to ADS. The outcome measures were mean differences in single longest look (single longest look to the target – single longest look to the non-target, in ms) in ADS, differences in single longest look in IDS, proportions of looking time to the target in ADS (%), proportions of looking time to the target in IDS (%), latency in ADS, and latency in IDS (ms) averaged across trials for each participant. Four participants were excluded from the regression analyses for 18 months and six were excluded for 24 months as their speech data were not usable for prosodic analyses. The descriptive statistics are shown in Table 3. Supplementary Figures 1–4 show scatter plot matrices of correlations (Pearson correlation coefficients) between all predictors (raw prosodic values and prosodic hyper-scores) and children's looking behaviors.

We first examined whether there were correlations between mothers' IDS prosody (raw prosodic values) and children's looking behavior in word learning. Regression analyses revealed that, for the 18-month-old group, the final model showed no significant predictors between the raw prosodic values and the differences in single longest look in the ADS condition. The final model for the IDS condition ($R^2 = 0.27$, $F(1, 18) = 6.57$, $p = 0.020$) showed that utterance pitch range was positively correlated with children's differences in single longest look ($\beta = 191.49$, $SE = 74.72$, $t = 2.563$, $p = 0.020$), suggesting that a larger pitch range in mothers' IDS at 18 months is correlated with better word learning from IDS at 24 months (Fig. 5). The final models for proportion of looking time and latency showed no significant predictors. For the 24-month-old group, none of the models showed significant predictors for children's word learning in either ADS or IDS.

We then examined whether there were correlations between prosodic hyper-scores and children's looking behavior in word learning. Regression analyses revealed no significant predictors in any of the models for either 18 or 24 months.

4. Discussion and conclusions

To understand the role of prosodic input in children's word learning, the current study had two goals. First, we conceptually replicated Ma et al. (2011) to investigate whether 24-month-old Dutch children learned novel words better from prototypical IDS than ADS. Second, we explored the link between individual mothers' prosodic input and children's word learning from ADS and IDS.

Regarding the first goal, our findings show that Dutch 24-month-old children successfully learned novel words from both ADS and IDS, yet there is evidence that IDS still facilitates children's word learning compared to ADS at this age. This finding was robust with respect to the measure of single longest look, the same measure used in Ma et al.'s original study. Compared to Ma et al. (2011), who found that American-English-learning 21-month-old children could only learn novel words from prototypical IDS but not ADS, our findings suggest that the prosody of IDS facilitates language acquisition even when children are efficient word learners at 24 months.

In addition to the measure of single longest look, we adopted two other measures of children's looking behavior: proportions of looking time and latency, which were also commonly used in analyzing data from the Intermodal Preferential Looking Paradigm (Delle Luche et al., 2015). The facilitative effect was only observed on the measure of single longest look, the same measure that was chosen by Ma et al. (2011). As the authors noted, "The dependent variable was single longest look at the target and the nontarget in each test trial, which has been found to be a more sensitive measure than total looking time, especially when children's attention decreases over the course of a test (e.g., Schafer & Plunkett, 1998; Bailey & Plunkett, 2002)." (Ma et al., 2011, p. 215). This was in accordance with Delle Luche et al. (2015), who compared single longest look and proportions of looking time and found that the effect sizes were larger for the single longest look than for the proportions of looking time.

To explain the inconsistency between the measures, it is possible that the single longest look was a more sensitive measure of children's looking behavior, at least for the paradigm used in Ma et al. (2011) and the current study. In our correlational analyses, we only found a significant correlation between mothers' IDS prosody and children's single longest look differences, while no significant correlation was found for the measure of proportions of looking time. An alternative explanation is that as Dutch 24-month-old children can already learn novel words from IDS, the facilitative effects of IDS are not as robust as in Ma et al. (2011), in which children could not reliably learn novel words from ADS. For the measure of latency, we could only include trials in which children initially looked at the distractor. As a result, we had a limited number of trials in the data analysis. In this case, our results were somewhat preliminary for this measure. Alternatively, it is possible that the gaze shift latency was no longer a robust measure for learning performance when children could reliably learn novel words from ADS and IDS.

One reviewer proposed alternative interpretations of our data. First, it is possible that children learned the novel words better in ADS than in IDS and they got bored faster in ADS in the testing phase. Second, our results may be interpreted as evidence that IDS maintained the children's attention to the target or their interest in looking at the target more than ADS. We would argue that our facilitative effect was supported by the significant interaction, which suggests that children's single longest look to the Target (compared to Distractor) was specifically longer in IDS condition. In other words, the differences in looking time to the Target and Distractor were larger in the IDS condition. If children got bored faster in ADS, or IDS only attracted children's attention, we would have found a significant main effect of Condition, showing that children's looking time in the ADS Condition was significantly shorter than IDS Condition. In our data analysis, we also checked whether children paid more attention in the IDS Condition than in the ADS condition in the entire testing phase and the results showed no significant differences.

Thus far, two questions remain to be answered. First, when do children start to reliably learn novel words from ADS? Our findings suggest that 24-month-old Dutch-learning children could reliably learn novel words from ADS. Ma et al. (2011) found that 21-month--old American-English-learning children only learned novel words from IDS. Also, they found that children who had a relatively larger vocabulary learned novel words in ADS. It appears that as children grow older and their vocabulary size increases, they gradually rely less on IDS to learn novel words. However, as Ma et al. (2011) and the current study tested different age groups, the exact age when

Table 3
Means and standard deviations (*SD*s) of raw prosodic values of IDS and hyper-scores in 18 ($N = 20$) and 24 months ($N = 18$).

Age	18 months				24 months			
Prosodic measures	Raw prosodic values		Hyper-scores		Raw prosodic values		Hyper-scores	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Articulation rate (syllables/s)	4.58	0.75	0.89	0.16	4.77	0.47	0.94	0.09
Mean F0 (st)	28.25	2.11	1.07	0.07	26.89	1.78	1.05	0.11
F0 range (st)	15.59	3.16	1.11	0.29	15.08	3.64	1.08	0.30

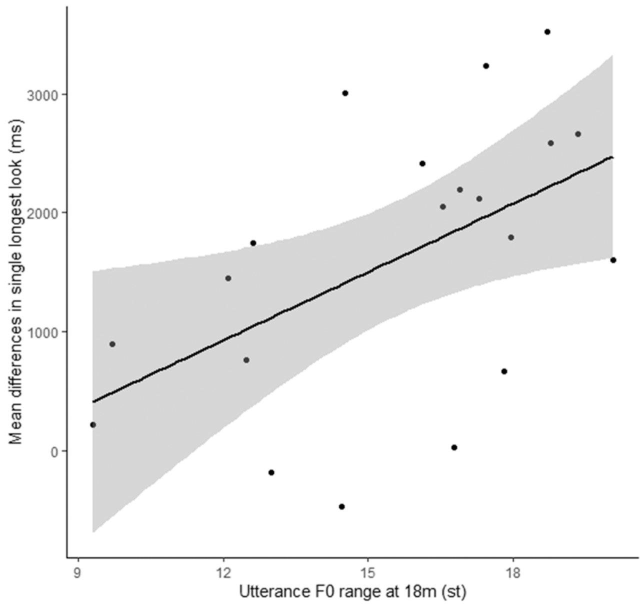

Fig. 5. Relationship between utterance F0 range at 18 m and mean differences in single longest look during word learning from IDS.

children no longer rely on IDS to learn novel words could not be concluded from these two studies. In addition, it remains a question whether children's trajectory of word learning from ADS and IDS varies among languages and cultures. As the degree of prosodic exaggeration seems to be most salient for American English IDS compared to other languages (e.g., Fernald, 1989), it is possible that American-English-learning children are used to the exaggerated prosody of IDS. Thus, it might take them longer to reliably learn novel words from ADS. On the other hand, for languages and cultures with a less exaggerated IDS, children may rely less on the prototypical IDS prosody to learn novel words and start to reliably learn novel words from ADS earlier. As we only tested Dutch children and did not directly compare them with other languages and cultures, we cannot draw any conclusions on this matter. So far only a handful of studies (including our own) have examined the effect of IDS on children's word learning experimentally and they focused on different populations and ages. Therefore, this question requires further empirical evidence from cross-linguistic and cross-cultural studies. Such comparisons could best be carried out in future multilab studies that control for data collection methods among languages. (e.g., ManyBabies Consortium, 2020).

Second, till when does the facilitative effect of prototypical IDS on children's word learning last? In a meta-analysis, Spinelli, Fasolo, and Mesman (2017) found that the association between prototypical IDS prosody and children's language outcomes were greater for pre-linguistic than linguistic outcomes. We found that Dutch IDS at 24 months still had distinctive prosody compared to Dutch ADS, including a higher pitch and a slower articulation rate. These results are in line with Han et al. (2020) on Dutch IDS. Also, the facilitative effect of IDS on language acquisition was still salient at 24 months. Such facilitative effects of IDS prosody may not be limited to child language acquisition. 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. Recent evidence suggests that adult native English speakers learn Chinese target words better when they are presented in IDS-like speech (Ma, Fiveash, Margulis, Behrend, & Thompson, 2020). As the authors pointed out, many properties of IDS, such as prosodic and phonetic salience, may assist learning by attracting more attention or enhancing memory. Based on these studies, IDS-like speech may be helpful for lexical processing even for adults. However, these studies differ from the developmental research in several ways. First, young children's primary input in their daily life was IDS and the individual differences in their prosodic input may influence their learning in ADS and IDS. Second, children's individual differences in their linguistic ability, for example, vocabulary size, may be associated with children's word learning performance (Ma et al., 2011). It is still important to further test whether the benefits of IDS change as children grow older and gain a larger vocabulary size during the first years of life.

We further explored the relationship between individual mothers' prosody and children's word learning from ADS and IDS. Here, we focus on the prosody of one specific speech context: word-learning contexts in which mothers introduce unfamiliar words to children. Our findings suggest that mothers' IDS pitch range in word-learning contexts is correlated with children's word learning from IDS. Mothers who had a larger pitch range at 18 months had children who performed better in word learning tasks when the stimuli had prototypical IDS prosody. On the other hand, there was no significant predictor for children's looking behavior in word learning from ADS. These results suggest that children's word learning from IDS is related to their prior experience with prosodic input in word-learning contexts, which is crucial for children to learn words. Even though pitch range expansion does not seem to be a general feature of Dutch IDS, individual differences in IDS pitch range is still a significant predictor of children's word learning from IDS. Previous research suggests that children's vocabulary starts to increase quickly and their word learning skills improve significantly at around 18 months (Bion et al., 2013; Goldfield & Reznick, 1990). Our findings indicate that mothers' IDS prosody at the beginning of vocabulary growth might contribute to this rapid word learning period. There are two possible mechanisms behind the link between IDS prosody and children's word learning. First, it is possible that mothers fine-tune their speech prosody to highlight unfamiliar words when addressing children, which supports children's word learning directly. Second, it is possible that a larger pitch range in IDS attracts children's attention (Kitamura & Burnham, 2003), which may lead to a better learning effect.

In conclusion, our findings suggest that even though 24-month-old Dutch children can successfully learn novel words from both ADS and IDS, IDS still has a facilitative effect on children's word learning at this age. Children's word learning from IDS, but not ADS, is related to their prior experience with prosodic input in word-learning contexts. These findings contribute to an understanding of the role of IDS prosody in language development, highlighting individual differences and contextual differences in IDS prosody across speakers and languages.

Funding

This work was partly supported by Shanghai Planning Project of Philosophy and Social Sciences (2020EYY001); Fundamental Research Funds for the Chinese Central Universities Research, East China Normal University (43800–20101-222054); Universiteit Utrecht.

CRediT authorship contribution statement

Mengru Han: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft. **Nivja H. de Jong:** Conceptualization, Methodology, Writing – review & editing. **René Kager:** Conceptualization, Methodology, Writing – review & editing.

Acknowledgments

We thank Chris van Run for scripting the experiment. We thank Suzanne Dekker, Maartje de Klerk, Karlijn Kouwer, and Yuchen Li, for their help with data collection and annotation. We are grateful to all the families who participated in this study. This work was partly supported by Shanghai Planning Project of Philosophy and Social Sciences (2020EYY001); Fundamental Research Funds for the Chinese Central Universities Research, East China Normal University (43800–20101-222054); Universiteit Utrecht. We thank the reviewers for the CogSci 2022 conference for their useful suggestions on an earlier version of this manuscript. We acknowledge the editor and the anonymous reviewers for their valuable comments and suggestions.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.infbeh.2022.101728.

References

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. https://doi.org/10.1016/j.jml.2012.11.001

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

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

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

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

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.

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

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., & 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., 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 (Eds.), *Developmental psycholinguistics: Online methods in children's language processing* (pp. 97–135). Amsterdam, The Netherlands: John Benjamins.

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

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

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

Han, M. (2019). *The Role of Prosodic Input in Word Learning: A Cross-linguistic Investigation of Dutch and Mandarin Chinese Infant-directed Speech*. Amsterdam: LOT Publications.

Han, M., de Jong, N., & 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*, 85–111. https://doi.org/10.1017/S0305000919000813

Han, M., de Jong, N. H., & Kager, R. (2021). Language Specificity of Infant-directed Speech: Speaking Rate and Word Position in Word-learning Contexts. *Language Learning and Development*, 1–20. https://doi.org/10.1080/15475441.2020.1855182

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 (Eds.), *Methods for assessing children's syntax* (pp. 105–124). Cambridge, MA: MIT Press.

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, 45*(5), 1035–1053. https://doi.org/10.1017/S0305000917000629

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

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

Ma, W., Fiveash, A., Margulis, E. H., Behrend, D., & Thompson, W. F. (2020). Song and infant-directed speech facilitate word learning. *Quarterly Journal of Experimental Psychology, 73*(7), 1036–1054.

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.

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

ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science, 3*(1), 24–52. https://doi.org/10.1177/2515245919900809

Outters, V., Schreiner, M. S., Behne, T., & Mani, N. (2020). Maternal input and infants' response to infant-directed speech. *Infancy, 25*(4), 478–499.

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/.

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

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. (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

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.

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

Van de Weijer, J. (1999). *Language Input for Word discovery(Doctoral dissertation)*. Nijmegen: Radboud University Nijmegen. https://doi.org/10.17617/2.2057670

Van Rooijen, R., Bekkers, E., & Junge, C. (2019). Beneficial effects of the mother's voice on infants' novel word learning. *Infancy, 24*(6), 838–856.

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

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