Maternal input and infants’ response to infant-directed speech

Vivien Outters¹,² | Melanie S. Schreiner¹,²,³,⁴ | Tanya Behne²,⁵ | Nivedita Mani¹,²

¹Department for Psychology of Language, University of Goettingen, Göttingen, Germany
²Leibniz ScienceCampus “Primate Cognition”, Göttingen, Germany
³Clinic for Cognitive Neurology, University of Leipzig, Leipzig, Germany
⁴Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
⁵Department of Developmental Psychology, University of Goettingen, Göttingen, Germany

Abstract
Caregivers typically use an exaggerated speech register known as infant-directed speech (IDS) in communication with infants. Infants prefer IDS over adult-directed speech (ADS) and IDS is functionally relevant in infant-directed communication. We examined interactions among maternal IDS quality, infants’ preference for IDS over ADS, and the functional relevance of IDS at 6 and 13 months. While 6-month-olds showed a preference for IDS over ADS, 13-month-olds did not. Differences in gaze following behavior triggered by speech register (IDS vs. ADS) were found in both age groups. The degree of infants’ preference for IDS (relative to ADS) was linked to the quality of maternal IDS infants were exposed to. No such relationship was found between gaze following behavior and maternal IDS quality and infants’ IDS preference. The results speak to a dynamic interaction between infants’ preference for different kinds of social signals and the social cues available to them.

1 INTRODUCTION

Infants and young children are typically addressed using a different speech style—referred to as infant-directed speech (hereafter, IDS)—compared with adults. IDS differs from the kind of speech...
adults use to communicate, that is, adult-directed speech (hereafter, ADS), with typically higher pitch, greater pitch range, and longer pauses between words in IDS compared with ADS (Fernald et al., 1989). Caregivers produce IDS spontaneously, with variation in the quality of IDS across individuals, development, languages, and cultures (Cristia, 2013; Soderstrom, 2007). Furthermore, the use of IDS in interactions with infants plays a crucial role in early language development. First, we summarize findings from studies examining the role of IDS in early language development, before outlining the questions under investigation in this study.

1.1 What is infant-directed speech?

IDS and ADS differ along prosodic, phonolexical, and syntactic dimensions (Cristia, 2013; Soderstrom, 2007). At the prosodic level, the primary difference between IDS and ADS is the higher pitch used in interactions with infants across different age groups (Cristia, 2013), with increased variability in pitch, lengthening of vowels, and pauses in IDS relative to ADS. At the phonolexical and syntactic levels, IDS is linguistically simpler, including shorter, more redundant utterances, more questions, and increased repetition of single words (Grieser & Kuhl, 1988; Soderstrom, 2007; Soderstrom, Blossom, Foygel, & Morgan, 2008). Vowels produced in IDS show more disparate formant structures than in ADS (Kuhl et al., 1997). IDS also plays an affective role in early language, such that IDS is typically associated with the expression of positive emotions (Singh, Morgan, & Best, 2002).

However, there is considerable variation in the acoustic characteristics of IDS and the use of IDS in communication with infants across different linguistic cultures (Farran, Lee, Yoo, & Oller, 2016; Fernald et al., 1989; Schieffelin & Ochs, 1983; Werker et al., 2015). Research also suggests variation in the characteristics of IDS across development: While mothers’ speech increases in amount and complexity in interactions with 3-month-olds compared with 1-month-olds (Henning, Striano, & Lieven, 2005; see also Kaye, 1980), the mean length of utterance decreases between early infancy and the second year of life (Murray, Johnson, & Peters, 1990; Sherrod, Friedman, Crawley, Drake, & Devieux, 1977; Stern, Speiker, Barnett, & MacKain, 1983). At the segmental level, mothers differently exaggerate acoustic contrasts in IDS across development with, for instance, reduced voice-onset-time overlap between voiced and voiceless stops, in speech to 15-month-olds, while mothers of 6- to 9-month-olds produce these contrasts similarly across IDS and ADS (Malsheen, 1980; see Bernstein Ratner, 1984; Cristia, 2010 for other differences). Overall, while caregivers spontaneously produce IDS in communication with infants, there is considerable variation at the individual, developmental, cultural, and linguistic level in the characteristics of maternal IDS that infants are exposed to.

1.2 IDS preference

Neonates and young infants prefer to listen to IDS compared with ADS (e.g., Cooper & Aslin, 1990; Pegg, Werker, & McLeod, 1992). This preference remains relatively stable between 4 and 8 months (Fernald, 1985; Fernald & Kuhl, 1987; Glenn & Cunningham, 1983; Kaplan, Goldstein, Huckey, & Cooper, 1995; Kaplan, Goldstein, Huckey, Owren, & Cooper, 1995; Pegg et al., 1992; Schachner & Hannon, 2011; Singh et al., 2002; Werker & McLeod, 1989; but see Hayashi, Tamekawa, & Kiritani, 2001 below) while findings with infants aged 9–18 months are mixed (e.g., Glenn & Cunningham, 1983; Hayashi et al., 2001; Newman & Hussain, 2006; Segal & Newman, 2015). There are further differences across studies based on the language being tested, the method employed, the speaker, and the type of speech used in the experiments (Cooper, Abraham, Berman, & Staska, 1997; Kaplan et al.,...

1.3 | The functional relevance of IDS

Infants also learn better from stimuli presented in IDS compared with ADS: Studies suggest that infants segment words better from fluent speech when the speech is more infant directed (Floccia et al., 2016; Schreiner & Mani, 2017; Singh, Nestor, Parikh, & Yull, 2009; Thiessen, Hill, & Saffran, 2005). Infants are able to segment words from ADS (see Mani & Pätzold, 2016; Schreiner, Altvater-Mackensen, & Mani, 2016), but they do so more easily (with fewer cues and less exposure) in IDS. Infants also learn words easier from IDS relative to ADS (Graf-Estes & Hurley, 2013; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011), and the use of IDS in communication with infants correlates with increased vocabulary scores later in life (Fisher & Tokura, 1996; Vosoughi, Roy, Frank, & Roy, 2010; Weisleder & Fernald, 2013). Differences in infants’ brain responses to IDS and ADS across development (Naoi et al., 2012; Zangl & Mills, 2006) similarly speak to the brain’s dynamic and flexible response to IDS across development.

An influence of IDS on language development is typically explained using one of two accounts (Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015). Infants may learn language better from IDS because IDS provides more information, for example, more exaggerated cues to the boundaries between words, while at the same time providing information reduced in complexity (e.g., Fernald, 1989; Fernald et al., 1989). Alternatively, IDS may be more attractive, salient, and relevant to the infant and draw attention to the input, thereby bolstering learning (e.g., Zangl & Mills, 2006). It may be difficult to adjudicate between these two approaches in the linguistic domain as the information to be attended to and learned is necessarily presented in IDS (including any potential hyperarticulation that may equally boost learning).

An examination of the influence of IDS on infants’ performance in a nonlinguistic task may better inform us of the extent to which IDS perceptually attracts infants’ attention to the current interaction and supports learning. Such evidence is provided by work examining the impact of IDS on joint attention (Senju & Csibra, 2008), showing that infants followed the gaze of a model more reliably and quickly when the interaction was cued with IDS relative to ADS provided, highlighting the functional relevance of IDS to interactions with infants (see also Gredebäck, Astor, & Fawcett, 2018).

1.4 | The current study

These findings reveal that (a) caregivers spontaneously produce IDS albeit with variation across caregivers, cultures, languages, and development, (b) infants show a preference for IDS over ADS which also varies across development, (c) IDS has a beneficial impact on language development, and (d) IDS is functionally relevant even in nonlinguistic tasks, highlighting the salience of IDS in interactions with young infants.

While the aspects of IDS in parent–infant communication highlighted above have been much studied in isolation, such studies do not allow an examination of the intriguing possibility that they may be related. Recent work highlights a positive relationship between maternal IDS prosody and later linguistic outcomes and infant attention (see Spinelli, Fasolo, & Mesman, 2017 for a review
and a recent meta-analysis). These studies suggest that pitch modulation of maternal IDS has a significant influence on infants’ later pre-linguistic and linguistic outcomes (e.g., D’Odorici & Jacob, 2006; Lyakso, Frolova, & Grigorev, 2014) including their later joint attention skills (Roberts et al., 2013). Equally, studies suggest that infants respond better in the moment to stimuli that are more prosodically exaggerated (e.g., Butler, O’Sullivan, Shah, & Berthier, 2014; Niwano & Sugai, 2002a, 2002b).

Tying these two veins of research together, this study examines whether there is a relationship between infants’ in-the-moment response to IDS and the prosodic properties of maternal IDS they have been exposed to. Such a relationship is presupposed in theories examining the role of IDS in learning. For instance, pedagogical theories of learning suggest that parents employ IDS as an ostensive cue to the infant in order to signal that the interaction is relevant to them and tailor their input to the child in order to provide them with optimal input (Csibra & Gergely, 2009, 2011). Evolutionary developmental theories of language acquisition highlight the importance of IDS in capturing infants’ in-the-moment attention in day-to-day interaction, with parents being sensitive to the attentional salience of IDS and modulating their speech based on their infants’ responses (McMurray, 2016; Smith & Trainor, 2008). Thus, both these different theoretical accounts imply a close relationship between maternal IDS and infants’ response to IDS. Against this background, this study examines whether and to what extent such a relationship exists, that is, whether the quality of maternal IDS covaries with infants’ attention and response to IDS, in general.

We tested the same infants on their preference for listening to IDS, the functional relevance of IDS, and individual differences in the prosodic characteristics of IDS produced by their caregivers, to examine the interaction between these factors in parent–infant communication. First, we analyzed recordings of mothers producing IDS and ADS to examine whether the quality of German mothers’ speech differs across communication with infants and adults. We focused here on the suprasegmental characteristics of IDS given the studies showing an influence of maternal prosody on infants’ later language outcomes. Second, a preferential listening task examined whether infants of those mothers prefer listening to IDS over ADS (Cooper et al., 1997). Third, a gaze following task explored infants’ gaze following after someone “speaking” IDS and the same person “speaking” ADS (Senju & Csibra, 2008).

We tested different groups of infants at two specific points, that is, at 6 and 13 months of age, to examine if these aspects of IDS in early life change across development. We target these age groups specifically given previous findings of differences in the neural and behavioral responses to IDS and ADS (Naoi et al., 2012; Schreiner & Mani, 2017; Zangl & Mills, 2006), as well as differences in the influence of maternal IDS on learning in younger and older infants (see Spinelli et al., 2017 for meta-analysis). Examining the interaction between these findings at two different ages allows us to investigate how these factors develop across infancy. Given the extensive testing schedule, involving three tasks spread across 2 days (in each age-group), we collected data from two separate groups of infants at 6 and 13 months.

We expect German mothers to produce systematic prosodic differences in their IDS and ADS at both ages tested (Fernald & Simon, 1984; Fernald et al., 1989). It is less clear whether these differences will vary across the two ages, given the limited research on prosodic differences in IDS across development. Based on previous research, we expect 6-month-olds to show a preference for listening to IDS compared with ADS, while it is less clear whether this preference will remain at 13 months of age (e.g., Newman & Hussain, 2006). Our hypotheses for the gaze following task are similar: 6-month-olds ought to follow the gaze of someone faster in the IDS relative to the ADS condition (Senju & Csibra, 2008) while it remains uncertain how the 13-month-olds might respond. Studies suggest a developmental change in infants’ gaze following behavior between 6 and 13 months (e.g.,
The interaction between the individual tasks allows us to examine (a) the potential factors underlying infants’ IDS preference by investigating the extent to which this preference is related to the kind of language input infants are exposed to; (b) the extent to which a preference for listening to IDS is functionally relevant in nonlinguistic tasks, such as the gaze following task; and (c) the extent to which this functional relevance of IDS may be related to the kind of input available to the infant. Taken together, the simultaneous analysis of the three tasks in the same group of infants allows us to examine potential variation in interaction between these aspects of IDS processing across development.

2 | METHOD

2.1 | Participants

Eighty-one monolingual German infants were recruited. Forty-two infants at the age of 6 months (5 m 16 d – 7 m 10 d, 19 girls) and 39 infants at the age of 13 months (12 m 15 d – 14 m 12 d, 19 girls) took part in the study. Participants were excluded due to illness on the second day of testing (n = 8), fussiness on the part of the baby (n = 19), experimental error (n = 3), and data loss (n = 10) (for a list of participants see Table 1).

2.2 | General procedure

The study consisted of three subtasks: a preferential listening task, a gaze following task, and recordings of maternal speech. The tasks were split across 2 days. On the first day, infants participated in the preferential listening task, in one of the conditions of the gaze following task, and maternal ADS was recorded. After 5–30 days, infants returned to take part in the second condition of the gaze following task and the mother was recorded producing IDS. Additionally, infants also took part in an EEG experiment, whose results are not reported in this study. This study was conducted according to the guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the Ethics committee of the University of Goettingen's Psychology Institute (Number 153).
2.3 | Recordings

2.3.1 | Procedure

Mothers were recorded on two separate days, producing ADS and IDS. At the first appointment, the mother watched a short video (1 min) about a teddy bear having a small adventure, including different locations and people in order to retell the story shortly after. When the mother felt confident about retelling the story, she walked into a room, sat on a chair facing a woman’s face displayed on screen. She was told to imagine retelling the story to a friend. Recordings were obtained via a microphone placed in front of the mothers’ feet. While this may have led to differences in the quality of maternal ADS due to parents being left on their own rather than in interaction with another adult, we note that the properties of maternal ADS recordings were similar to those reported in the literature (Fernald et al., 1989). We did not ask the parent to tell the story to another adult in case they felt self-conscious being recorded speaking to an unfamiliar adult, thereby impacting the quality of their speech.

At the end of the second appointment, mothers were recorded telling the same story in IDS. The mother could watch the video again in order to retell the story to her own infant. The experimenter left the room but recorded the interaction between mother and infant via a microphone and a camera placed about 1.5 m away. We did not counterbalance the order of recordings to ensure that the quality of ADS recordings was not influenced by the previous production of similar sentences in IDS. We assumed that the prior production of the stimuli in IDS would be more likely to influence the quality of maternal ADS (based on our experience recording stimuli across numerous studies) than the other way around.

2.3.2 | Data analysis

Average duration of IDS recordings was 152 s with an average of 84 s analyzable speech (minimum: 28 s, maximum: 210 s). ADS recordings were on average 118 s with an average duration of 56 s analyzable speech (minimum: 25 s, maximum: 126 s). We calculated the mean, maximum, minimum pitch, pitch range, and average duration of utterances separately for the IDS and ADS recordings using PRAAT software (Boersma & Weenink, 2019). Mothers provided, on average, 22 noise-free utterances in IDS (SD = 11.1) and 19 noise-free utterances in ADS (SD = 8.9). Paired samples t tests examined whether there was a difference in the prosodic characteristics of maternal IDS and ADS at both age groups. We also calculated the difference in the prosodic characteristics of maternal IDS and ADS and then compared this difference across the two age groups.

2.4 | Preferential listening task

2.4.1 | Material and design

A female native speaker of German recorded six short stories for this task (see Appendix 1). Each story was recorded twice, once in IDS and once in ADS. Acoustic analysis using PRAAT software (Boersma & Weenink, 2019) ensured the validity of the stimuli. Paired samples t tests run across individual sentences in each story revealed significant differences in the maximum pitch, t(41) = 15.06, p < .001, minimum pitch, t(41) = 6.17, p < .001, and mean pitch, t(41) = 15.14, p < .001, of IDS and ADS stimuli. There were no significant differences in the pitch range, t(41) = 0.41, p = .681, and
duration, \( t(41) = 1.310, p = .198 \) (see Table A1: Appendix 2 for a list of characteristics). The acoustic characteristics of the stimuli were similar to previous reports of naturalistic German IDS and ADS (Fernald et al., 1989).

### 2.4.2 Procedure

Each infant was tested in a quiet room. Infants sat on the parent’s lap or in a car seat about 100 cm away from a monitor. This monitor presented infants with a blinking checkerboard. The auditory stimuli were presented via loudspeakers, located above the television screen. Two cameras directly above the screen recorded the infant’s eye movements. Synchronized signals from the two cameras were routed via a digital splitter to provide two separate time-locked images of the infant, which were used for online and offline coding.

Stimuli were presented using the Look software (Meints & Woodford, 2008). During each trial, infants were presented with a blinking checkerboard on the screen. Each trial began with a visual and auditory attention getter to direct the infant’s attention to the screen. The auditory stimulus was presented once the infant fixated the screen. An experimenter, seated in an adjacent room watching a video of the infant, indicated whether the infant was looking at the screen or away by pressing a button on a keyboard. The screen remained black in between trials until the attention getter started the next trial. Infants listened to alternating IDS and ADS trials each with a different story (with a maximum length of 30 s). Infants listened to six trials, half of which were presented in IDS and half of which were presented in ADS—counterbalanced across infants for the register. Half of the participants were presented with an IDS story in the first and an ADS story in the second trial, and vice versa. Subsequent trials alternated between conditions (IDS-ADS-IDS-ADS-IDS-ADS or ADS-IDS-ADS-IDS-ADS-IDS). The trial continued until the end of the auditory stimulus or until the infant looked away from the screen for more than 2 s. A second independent coder evaluated a random sample of 10% of each condition offline. The inter-rater reliability was high (\( r = .937 \)).

### 2.4.3 Data analysis

All results refer to the mean duration of the first fixation at the checkerboard averaged for each participant across trials separated by condition. We only included data from infants who provided data for both conditions in the first half of the experiment, that is, infants needed to provide data for at least one trial in each condition in the first three trials of the experiment. As looking times drop dramatically across the duration of this task, this criterion avoided differences across conditions, which were driven by the drop in infants’ interest toward the latter half of the experiment. This led to exclusion of two infants in each age group.

Data were analyzed using linear mixed models, ANOVAs, and \( t \) tests. Due to the consistency across these analyses, we report the results of performance within each task using the ANOVAs and \( t \) tests and only report the results of the linear mixed-effects models when reporting the correlations across task. The dependent variable was the listening time to IDS and ADS trials. We ran a mixed ANOVA with the within-subjects factor condition (IDS, ADS), and between-subjects factors age (6, 13 m) and order of condition (IDS first, ADS first). The order of presentation of trials was included based on results in the literature (Cooper et al., 1997; Kaplan et al., 1995; but see also Cooper & Aslin, 1990, 1994) suggesting that preferential listening data are strongly influenced by the order of presentation of trials in the study with infants attending more to the early trials in the experiment than the later trials.
2.4.4 | Performance across task

We examined the correlations in performance across the three tasks (recordings of maternal input, preferential listening, and gaze following) using linear mixed-effects models. These models offer the advantage that we can control for performance in one task while examining the relationship in performance between two other tasks.

For the preferential listening task, we ran a linear mixed-effects model including condition (IDS, ADS), age (6, 13 m), and order of presentation (IDS first, ADS first) as fixed effects, and subject and item as random effects. We compared this model with a reduced model, excluding condition, to examine the effect of condition on the response. To investigate the performance across tasks, we included variables capturing infants' performance in the gaze following task and maternal IDS quality. Specifically, we included the difference in proportion of target fixations in the IDS and ADS conditions in the gaze following task and the difference in critical prosodic features of maternal IDS and ADS. Thus, the model included the interaction between condition and the prosodic characteristics of maternal IDS, for example, the difference among maternal mean pitch, pitch range, and utterance duration in IDS and ADS and the interaction between condition and the difference in the proportion of target fixations in the IDS and ADS condition in the gaze following task, with subjects, order of presentation, and item as random effects (where possible without violating singularity of fit).

A drop1 analysis then examined the key factors that, in interaction with condition (IDS and ADS trials in the preferential listening trials), explained the variance in listening times. This analysis circulated between removing specific variables from the model to estimate the best model fit. We only considered factors that influenced listening times in interaction with condition as main effects would merely indicate that a particular factor impacted listening times overall, and not suggest an effect of this factor on infants' IDS preference.

Finally, following the suggestion of one of the reviewers, we calculated the difference in the acoustic characteristics of maternal IDS and ADS and the acoustic characteristics of IDS and ADS stimuli presented in the preferential listening task (\((\text{IDS}_{\text{pref}} - \text{ADS}_{\text{pref}}) - (\text{IDS}_{\text{mat}} - \text{ADS}_{\text{mat}})\)) and examined whether adding this difference (using a similar drop1 analysis) improved model fit of the preferential listening time data. This analysis allows us to examine how infants' preference for the IDS presented at test varied as a factor of how similar this IDS was to the IDS they received at home.

2.5 | Gaze following

2.5.1 | Material and design

The visual stimuli used for the gaze following task were taken from Senju and Csibra (2008). Each video started with a woman, sitting behind a table in the middle of the screen, facing down, with two objects on the table positioned to the left and right side of the model. The objects remained on the table in the same location for the entire duration of the trial. The first part of the video presented infants with the model looking down between the two objects for 2 s. Then, an icon appeared on the head of the woman accompanied by a voice saying “Hallo” (“Hello”). The “Hallo” was uttered by a female German speaker in IDS or in ADS. The acoustic characteristics of the auditory stimuli were similar to those presented in Senju and Csibra (2008), see Table A2: Appendix 2. During the critical part of the trial, the icon disappeared and the model turned her head and shifted her gaze to one of the two objects on the table (1 s) and continued to fixate this object for the remaining 5 s.
2.5.2 | Procedure

Each infant was tested in a quiet room. Infants either sat on the parent's lap or in a car seat about 65 cm away from a monitor. Each infant was presented with six trials on each day of testing. Each trial contained a unique pair of objects. The direction of the model’s gaze was, as in Senju and Csibra (2008), counterbalanced in ABBABA order. For half of the infants, the model fixated the object to her right in the first trial, while she fixated the object to the left for the other half of the infants. After each trial, a jumping teletubby appeared to reorient the infants’ attention to the middle of the screen. In contrast to Senju and Csibra (2008), we used a within-subject design, where each infant was presented with the IDS condition on 1 day and the ADS condition on another. The assignment of infants to the IDS or ADS condition on the different days was counterbalanced.

2.5.3 | Data analysis

Areas of interest were defined based on the size of the objects on the table (340 × 340 pixel) and icon on the head of the woman (280 × 278 pixel). An automated eye tracker, Tobii X120, provided an estimate of where infants were looking with one data point every 8 ms. Data were included and further analyzed only if the eye tracker recorded reliable data from one of the eyes (validity less than 2 on the Tobii scale). Gaze data were further aggregated into 40 ms bins. Each of these 40 ms bins was coded for whether the infant was looking at the model, the target, or the distractor. These bins were aggregated according to our time window of interest. Our critical time window began at 4,000 ms with the offset of the icon above the model until the end of the trial (10,000 ms). We excluded trials where infants did not fixate the model for at least 500 ms during the speech presentation, to ensure we only included trials where infants were attending to the model and her subsequent eye movement. This criterion excluded 8% of all trials.

We analyzed the proportion of target looking (duration of looking to target/looking to target, distractor, and model) in the critical phase of the trial, that is, after the model had shifted her gaze and fixated one of the objects, using a repeated-measures ANOVA with the within-subjects factor condition (IDS, ADS), and the between-subjects factors age (6, 13 m) and order of presentation (IDS first, ADS first).

2.5.4 | Performance across task

To investigate the influence of maternal input and preferential listening on the gaze following behavior, we run a linear mixed-effects analysis including condition (IDS, ADS), and age (6, 13 m) as fixed effects, and subject and item as random effects. We included variables capturing the quality of maternal IDS and infants’ performance in the preferential listening task. Performance in the preferential listening task was indexed by the difference in infants’ listening times to IDS and ADS. Maternal IDS quality was indexed using the difference among maternal mean pitch, pitch range, and utterance duration in IDS and ADS. Drop1 analyses examined the key factors that, in interaction with condition

---

1Following the conventions employed in our previous work, we code a child as looking at the target in a particular 40 ms bin if the child looked at the target for more than 2/3rds of the number of bins included in any 40 ms bin. If the child looked less at the target than this threshold, the child was coded as not looking at the target or looking wherever the child did look for at least 2/3rds of the bin (e.g., Eiteljörge, Adam, Elsner, & Mani, 2019).
(IDS and ADS trials in the gaze following task), explained the variance in target fixations. As above, we only considered factors that influenced fixations in interaction with condition.

### RESULTS

#### 3.1 Recordings of mothers’ speech

We analyzed the acoustic characteristics of mothers’ speech in both IDS and ADS, with regard to her maximum pitch, minimum pitch, mean pitch (each in Hz), pitch range (in semitones), and the mean duration of utterances (in seconds per utterance, see Table 2). In all cases and at both age groups, the prosodic characteristics of maternal IDS differed significantly from the characteristics of maternal ADS (see Table 1). We also compared the difference in these acoustic characteristics in IDS and ADS, for example, maximum pitch in IDS—maximum pitch in ADS, across the two ages and found no differences in these acoustic characteristics across IDS and ADS between the two ages tested (see

<table>
<thead>
<tr>
<th>Acoustic characteristics—mothers of 6-months-olds</th>
<th>Comparing IDS and ADS in mothers of 6-month-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Max $f_0$</td>
<td>345</td>
</tr>
<tr>
<td>Min $f_0$</td>
<td>183</td>
</tr>
<tr>
<td>Mean $f_0$</td>
<td>256</td>
</tr>
<tr>
<td>Duration</td>
<td>2.27</td>
</tr>
<tr>
<td>$f_0$ range</td>
<td>11.62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acoustic characteristics—mothers of 13-months-olds</th>
<th>Comparing IDS and ADS in mothers of 13-month-olds</th>
<th>Comparing the relative IDS–ADS difference between mothers of 6- and 13-month-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Max $f_0$</td>
<td>353</td>
<td>45</td>
</tr>
<tr>
<td>Min $f_0$</td>
<td>176</td>
<td>34</td>
</tr>
<tr>
<td>Mean $f_0$</td>
<td>259</td>
<td>30</td>
</tr>
<tr>
<td>Duration</td>
<td>1.99</td>
<td>0.32</td>
</tr>
<tr>
<td>$f_0$ range</td>
<td>12.80</td>
<td>3.93</td>
</tr>
</tbody>
</table>

Note: Mean, maximum, and minimum $f_0$ are reported in Hz, duration in seconds, and $f_0$ range in semitones per second.
Table 2). Thus, mothers did not produce more exaggerated IDS (relative to their ADS) between the two age-groups tested.

3.2 | Preference

A mixed ANOVA with the dependent variable listening time, the within-subjects factor condition (IDS, ADS), and between-subjects factors age (6, 13 m) and order of presentation of condition (IDS first, ADS first) found a significant main effect of condition, $F(1, 69) = 4.16, p = .045, \eta^2_G = 0.01$, and a significant interaction among condition, age, and order of presentation of condition, $F(1, 69) = 5.27, p = .025, \eta^2_G = 0.01$, but no other main effects. Pivoting on age, we found that 6-month-olds listened longer to IDS compared with ADS trials, $F(1, 36) = 8.86, p = .005, \eta^2_G = 0.01$, with an interaction between condition and order of presentation of condition, $F(1, 36) = 17.98, p < .001, \eta^2_G = 0.04$. With 13-month-olds, we found no main effect of condition and only an interaction between condition and order of presentation of condition, $F(1, 33) = 25.89, p < .05, \eta^2_G = 0.16$. Splitting these interactions at both age groups, we found that 6-month-olds who received the IDS trial first in the experiment showed an IDS preference, $t(18) = 4.48, p < .001, 95\%$ CI: 2.06, 5.7, while those who received an ADS trial first did not, $t(18) = −1.06, p = .301, 95\%$ CI: $−2.02, 0.67$. Thirteen-month-olds showed a preference for ADS when the experiment began with an ADS trial, $t(17) = −3.82, p = .001, 95\%$ CI: $−7.39, −2.13$, and a preference for IDS when the experiment started with an IDS trial $t(16) = 3.39, p = .004, 95\%$ CI: $1.72, 7.49$. Figure 1 plots the mean listening times to each condition separated by age and condition presented first.

3.2.1 | Correlation across tasks

We ran a linear mixed-effects model on the data from the preferential listening task, that is, listening times, with the interaction between condition (IDS, ADS) and the acoustic characteristics of the difference in maternal IDS and ADS (i.e., mean pitch difference, difference in pitch range, and utterance duration difference) and the difference in target looking across the IDS and ADS...
condition in the gaze following task (see below for further details on this measure) as fixed effects and random effects on subject, order of presentation, and item (where possible given problems of singular fit when controlling for variance at the item level). Drop1 analysis found a significant interaction between condition and mean pitch difference at 6 months of age, $\chi^2(1) = 5.34, p = .021$, and an interaction between condition and duration of utterances at 13 months of age, $\chi^2(1) = 6.99, p = .008$ (see Table 2 for estimates). This pattern was corroborated by correlation tests between listening times in the IDS condition and the mean pitch difference ($r = 0.48, p = .026$) between maternal IDS and ADS. Note, this pattern changed when considering the total duration of listening time to IDS and ADS trials in the preferential listening task: Drop1 analysis found no significant interaction between condition and mean pitch difference at 6 months of age (although a larger model including maximum and minimum pitch did find an interaction between pitch difference and condition, $\chi^2(1) = 6.32, p = .011^2$), but an interaction between condition and duration of

2We initially ran larger models including maximum, minimum, mean pitch difference, pitch range and utterance duration. We do not report the results of these larger models given potential issues of collinearity across these measures.

**TABLE 3** Model examining listening times in the preferential listening task with interaction between condition (IDS, ADS) and difference in mean pitch, duration, pitch range across maternal IDS and ADS; as well as difference in fixations to the target in IDS and ADS trials in the gaze following task as fixed effects and random intercept on subject split by age (6, 13 m): lmer(listening times ~ condition * (mean pitch + duration + pitch range + PTL) + (1 | first) + (1 | subject).

<table>
<thead>
<tr>
<th>Predictors</th>
<th>6-month-olds</th>
<th>13-month-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimates</td>
<td>CI</td>
</tr>
<tr>
<td>Intercept</td>
<td>5.00</td>
<td>−0.44 to 10.43</td>
</tr>
<tr>
<td>Condition:mean pitch</td>
<td>0.05</td>
<td>0.01 to 0.10</td>
</tr>
<tr>
<td>Condition:duration</td>
<td>−0.38</td>
<td>−4.44 to 3.68</td>
</tr>
<tr>
<td>Condition:pitch range</td>
<td>0.26</td>
<td>−0.05 to 0.56</td>
</tr>
<tr>
<td>Condition:target fixations</td>
<td>−2.27</td>
<td>−18.22 to 13.68</td>
</tr>
</tbody>
</table>

*Note: Bold values highlight significant p-values.*
utterances at 13 months of age, $\chi^2(1) = 8.29, p = .003$. While the data entered into the model were continuous, for ease of depiction and interpretation, Figure 2 presents the averaged data for the interaction between condition mean pitch (2A) at 6 months of age and the interaction between condition and duration at 13 months of age. The estimates given in Table 3 provide a picture of the variance explained by these factors at each age group.

We also calculated the difference between the acoustic characteristics of IDS and ADS of the maternal recordings and the acoustic characteristics of IDS and ADS stimuli presented in the preferential listening task ($\text{IDS}_{\text{pref}} - \text{ADS}_{\text{pref}} - (\text{IDS}_{\text{mat}} - \text{ADS}_{\text{mat}})$). The greater this difference was, the more exaggerated was the difference between IDS and ADS stimuli presented in the preferential listening task relative to the difference between the mothers’ own IDS and ADS. A drop1 analysis found that adding this difference improved model fit at 6 months of the preferential listening time data with regards to mean pitch, $\chi^2(1) = 5.77, p = .016$. At 13 months we found improved model fit with regards to duration $\chi^2(1) = 5.12, p = .023$. In particular, infants showed a familiarity preference at 6 months, with infants, whose mothers similarly exaggerated IDS relative to ADS as was the case in the stimuli presented in the preferential listening task, showing a greater IDS preference. In simplified terms, infants preferred listening to IDS that was more similar to their mothers’ IDS.
3.3 Gaze following

Figure 3 plots the proportion of target looking after the model had initiated her gaze to the target object separated by condition throughout trials (IDS and ADS). We analyzed the proportion of looking to the target after the model had shifted her gaze and fixated one of the objects presented on the table in front of her. A repeated-measures ANOVA with the within-subjects factor condition (IDS, ADS), and the between-subjects factor age (6, 13 m) and order of condition (IDS first, ADS first) revealed a significant main effect of condition, \( F(1, 57) = 7.31, p = .009, \eta^2 = 0.03, \) indicating increased target looking in IDS compared with ADS (see Figure 4, with a split by age group). No other main effects or interactions were found. These results suggest no differences across age groups with regard to the influence of IDS on infants’ gaze following behavior.

3.3.1 Correlation across tasks

We ran a linear mixed-effects model on the fixations to the target, with the interaction between condition (IDS, ADS) and the acoustic characteristics of the difference in maternal IDS and ADS (i.e., mean pitch difference, utterance duration, and difference in pitch range) and the normalized difference in listening times to IDS and ADS trials in the preferential listening task as fixed effects and random effects on subject and item. There were no significant interactions between condition and any of the above factors in this analysis, \( ps > .112. \)

4 DISCUSSION

The aim of this study was to investigate the interaction among the prosodic characteristics of maternal IDS, infants’ preference for IDS over ADS, and the functional relevance of IDS in infant-directed communication, that is, the extent to which IDS impacts infants’ responding. We hypothesized that these three facets of IDS may influence one another, such that variation in one of these factors co-occurs with variation in the other. In particular, we were interested in (a) whether there are prosodic differences in the IDS and ADS produced by German mothers, (b) whether German infants show a preference for IDS compared with ADS, (c) whether German infants follow someone’s gaze faster when this person was previously presented speaking IDS compared with ADS, (d) how these characteristics vary across development, and (e) if and how these factors interact with one other. We were interested in the potential roots of infants’ preference for IDS and the extent to which this may be related to the quality of maternal IDS, and the extent to which infants’ preference for IDS may itself be related to the functional relevance of IDS in communication.

Similar to other studies (e.g., Fernald et al., 1989), there were significant differences in the prosodic characteristics of German mothers’ IDS compared with their ADS at both age groups. Mothers of infants at 6 and 13 months produced utterances with higher mean, maximum, and minimum pitch in IDS compared with ADS. They produced utterances of shorter duration in IDS relative to ADS. While utterances in IDS tend to be longer due to the lengthening of vowels and longer pauses, the total duration of utterances in naturalistic IDS tends to be shorter due to the fewer number of words per utterance (see Soderstrom, 2007, for a review). They also produced utterances with a narrower pitch range in IDS compared with ADS. This may be because mothers produced utterances with higher minimum pitch in IDS, thereby potentially reducing the range possible in this register.
We found no developmental differences in IDS produced by mothers of 6- and 13-month-olds. In other words, 13-month-olds hear IDS which is prosodically similar to that heard by 6-month-olds. Admittedly, these findings are in the context of mothers’ retelling a story to their infant, which may differ from the prosodic characteristics of their other interactions with their infant. However, ongoing data from our laboratory examining natural interactions between mothers and infants at the same ages similarly find no differences in the prosodic characteristics of maternal IDS across early infancy (c.f., Schreiner, Eiteljoerge, Behne, & Mani, in prep).

A preference for IDS over ADS was only found at 6 months of age. This finding is in line with previous studies reporting an IDS preference between 4 and 8 months (Cooper et al., 1997; Fernald, 1985; Kaplan, Goldstein, Huckeby, & Cooper, 1995; Kaplan, Goldstein, Huckeby, Owren, et al., 1995; Newman & Hussain, 2006; Singh et al., 2002; Thiessen et al., 2005; Werker & McLeod, 1989). At 13 months, we did not find a preference for IDS over ADS. This is likely attributable to changes in the speed of infants’ language processing from 6 to 13 months of age, with older infants being able to process a broader range of stimuli (c.f., Schreiner & Mani, 2017), including less infant-directed stimuli, leading to 13-month-olds not showing as much as a preference for IDS. Similar findings are reported in Newman and Hussain (2006).

These findings contrast with findings of a decrease in preference between 7 and 10 months, and a later increase in preference after this age (Hayashi et al., 2001). These also contrast with the findings from the ManyBabies project on IDS preference, with a more robust IDS preference past 12 months of age (The ManyBabies Consortium, 2020). This preference was more robust in North American English (NAE) infants relative to infants from non-NAE backgrounds and was tested using NAE IDS, which differs starkly from German IDS. We presented infants with German IDS and ADS, with the possibility that the reduced prosodic differences between German IDS and ADS (Fernald et al., 1989; Schreiner & Mani, 2017) lead to a reduced IDS preference at this later age.

In line with previous findings (Cooper et al., 1997; Kaplan, Goldstein, Huckeby, & Cooper, 1995; The ManyBabies Consortium, 2020; see also Cooper & Aslin, 1990, 1994), we found that the difference in listening times varied with order of presentation of trials. This was particularly so with the 13-month-olds who listened longer to whichever register they were presented with first. The 6-month-olds did not show increased listening times to ADS trials even when the first trial presented was in ADS. With all the constraints associated with interpreting null effects, this suggests conflicting influences from the order of presentation of trials and a potentially “real” IDS preference at 6 months. This is in line with our results, showing an overall IDS preference only at 6 months.

Despite the effect appearing smaller in the older age group relative to the younger age group, we do not find statistical differences in the performance of the gaze following task across the two age groups. Here, we found a significant effect of condition that did not differ across age groups, with 6- and 13-month-olds fixating the target more when cued by IDS than by ADS. This finding replicates Senju and Csibra’s (2008) findings with 6-month-olds (see also Hernik & Broesch, 2019) and extends this IDS effect to 13-month-olds, speaking overall to the functional relevance of IDS in communication with young infants. IDS appears to attract infants’ attention more and drives them to engage with others in social interactions; IDS is attentionally salient. As the task performed by infants was non-linguistic—and could not therefore be influenced by hyperarticulation in IDS—the findings highlight the attentional salience of IDS in driving infant responding and learning.

4.1 Performance across tasks

Before discussing these results, we highlight two features of this analysis which are key to interpretation. First, we only took the difference in the acoustic characteristics of individual mothers’ IDS and
ADS into the analyses, allowing us to tap into the infant directedness of maternal input and not the prosodic characteristics of the mothers’ speech, per se. Second, we controlled for variance in the other tasks when examining performance in any one task, that is, in analyzing the contribution of variation in maternal input on performance in the preferential listening task, we controlled for performance in the gaze following task, and vice versa.

Overall, we found that the acoustic characteristics of maternal IDS (relative to ADS) influenced performance most. Differences in the mean pitch and duration of utterances in maternal IDS (relative to ADS) influenced performance in the preferential listening task. At 6 months, infants whose mothers produced more exaggerated prosody (with regard to mean pitch) in IDS (relative to ADS) showed a greater preference for IDS (relative to ADS) in the preferential listening task. This finding highlights a keen link between the prosodic characteristics of maternal IDS and infants’ IDS preference. Examining the difference between maternal IDS and the IDS stimuli presented in the preferential listening task, we found that infants whose mothers’ IDS was similar to the IDS in the preferential listening task showed a more robust IDS preference, suggesting a familiarity preference underlying the IDS effect reported here. This raises questions as to the extent to which a similar familiarity preference underlies the IDS preference in the studies reported to-date, especially against the background of studies suggesting that infants show improved segmentation of words from exaggerated IDS even when the IDS typical to their native language is not similarly exaggerated (Floccia et al., 2016; Schreiner & Mani, 2017). Given the exploratory nature of these analyses, we do not comment on this issue further but highlight that different factors may underlie the reported IDS preference, including but not limited to the quality of the IDS that children are exposed to, and the quality of the stimuli presented to infants in a particular task.

At 13 months, we found that infants whose mothers produced shorter utterances in IDS (relative to ADS) showed a greater preference for IDS (relative to ADS) in the preferential listening task. Shorter utterances have been highlighted as one of the key features of maternal IDS and appear to interact with infants’ IDS preference at 13 months, as was the case with the pitch of maternal IDS at 6 months. We raise, however, the concern that the results with 13-month-olds need to be treated with considerable caution, given the presence of outliers that may have driven this effect and the fact that there was no independent IDS preference found at this age—we only found an interaction between condition and order of presentation at this age.

We found no influence of performance in the gaze following task on performance in the preferential listening task and vice versa. One reason why we found no influence of maternal IDS quality on gaze following performance may be due to the increased social cues provided in this task. The gaze cue, in conjunction with the increased attentional salience of IDS, may leave little room for modulation of infants’ performance by maternal IDS quality, due to the relevance of these cues. In other words, the increased social cues in the gaze following task may outweigh individual variation caused by infants’ preference for IDS (over ADS) or the acoustic characteristics of maternal IDS (relative to ADS). This explanation highlights the dynamic interaction between multiple cues in infant responding to IDS, with differential weighting of social cues, maternal input, and IDS salience in the same infants in different tasks across early language development.

5 | CONCLUSION

Our findings highlight developmental differences in infants’ preference for IDS over ADS, with 6-month-olds showing a preference while 13-month-olds do not. We do not find developmental differences in the functional relevance of IDS across development, or any in the quality of maternal IDS.
Across tasks, the findings highlight the strong link between infants’ response to IDS and the quality of maternal IDS, which may have consequences for our understanding of infants’ learning from IDS. We note that the direction of this effect remains unclear due to the cross-sectional nature of the study. In other words, the link between maternal IDS and infant responding can either be explained by mothers being conditioned to produce more exaggerated IDS when their infants respond more actively to this register (c.f., Smith & Trainor, 2008), or infants responding more actively to this register because they have more exposure to exaggerated IDS in their daily input. At present, we cannot adjudicate between these two possibilities. We suggest that the repeated findings of a link between maternal input and infants’ response to IDS in this study highlight the shared underpinnings of maternal IDS and infants’ response to IDS. This study, therefore, speaks to a dynamic interaction between infants’ preference for different kinds of social signals and the social cues available to them.

ACKNOWLEDGMENTS
This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project number 254142454/GRK 2070 and the Leibniz ScienceCampus Primate Cognition.

CONFLICTS OF INTEREST
The authors declare no conflicts of interest with regard to the funding source for this study.

ORCID
Vivien Outters https://orcid.org/0000-0002-2061-8583

REFERENCES


---

**How to cite this article:** Outters V, Schreiner MS, Behne T, Mani N. Maternal input and infants’ response to infant-directed speech. *Infancy*. 2020;00:1–22. https://doi.org/10.1111/infa.12334
APPENDIX 1

Stimuli recorded for preferential listening task

Short stories taken from: http://www.kidsnet.at/Deutsch/erlebnisbaum.htm


APPENDIX 2

Acoustic characteristic of IDS and ADS stimuli in preferential listening task and gaze following

TABLE A1 Acoustic characteristics of infant-directed speech (IDS) and adult-directed speech (ADS) stimuli in preferential listening task

<table>
<thead>
<tr>
<th>Preferential listening task</th>
<th>IDS</th>
<th>ADS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max pitch (Hz)</td>
<td>404.81</td>
<td>268.60</td>
</tr>
<tr>
<td>Min pitch (Hz)</td>
<td>139.77</td>
<td>102.38</td>
</tr>
<tr>
<td>Mean pitch (Hz)</td>
<td>241.66</td>
<td>195.31</td>
</tr>
<tr>
<td>Pitch range (st)</td>
<td>18.73</td>
<td>18.32</td>
</tr>
<tr>
<td>Duration (s)</td>
<td>2.92</td>
<td>3.30</td>
</tr>
</tbody>
</table>

TABLE A2 Acoustic characteristics of infant-directed speech (IDS) and adult-directed speech (ADS) stimuli in gaze following task

<table>
<thead>
<tr>
<th>Gaze following task (adapted German version and original English version)</th>
<th>German IDS</th>
<th>German ADS</th>
<th>English IDS</th>
<th>English ADS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max pitch (Hz)</td>
<td>494.12</td>
<td>291.56</td>
<td>538.56</td>
<td>272.14</td>
</tr>
<tr>
<td>Min pitch (Hz)</td>
<td>256.21</td>
<td>123.54</td>
<td>259.82</td>
<td>181.72</td>
</tr>
<tr>
<td>Mean pitch (Hz)</td>
<td>401.72</td>
<td>248.63</td>
<td>434.76</td>
<td>220.27</td>
</tr>
<tr>
<td>Pitch range (st)</td>
<td>11.37</td>
<td>14.86</td>
<td>12.61</td>
<td>6.98</td>
</tr>
</tbody>
</table>