Dear editor and reviewers,

Thank you for your detailed review of the manuscript. We have taken all the comments and suggestions into account and addressed them in the revised manuscript. First we have included additional recent literature in the introduction based on the reviewers' suggestions. Second, we have addressed comments pertaining to limited measurement reliability and taken steps to increase measurement reliability of the current study. In particular, we have decided to perform a full-null model comparison to test hypotheses. To maximise the sample size while taking into consideration limited reliability of the toddlers' preference for IDS, when testing H2a and H2b, we will weigh the contribution of the data point to the model by the number of completed trials per child and register, meaning that infants who have completed more trials will be weighted more heavily in the analysis. This will allow us to maintain a large sample while taking into consideration the limited reliability when participants don't complete the full experiment. Third, we have included questions on paternal attitudes and knowledge on language development, as well as questions regarding reading as an activity in order to distinguish between paternal caregiving and experience with reading, both of which will be used in the exploratory analyses to generate novel hypotheses for future studies. Finally, we have corrected some errors that were spotted in the first version of the manuscript, regarding trial numbers and type of auditory stimuli in the eye tracking task. Please see the following point-by-point responses to the reviewer's comments and concerns, as well as the highlighted text in the manuscript for edits and added text.

We believe that we have addressed the concerns and issues raised by the reviewers and that it has resulted in an improved manuscript. We would like to thank you for your time and feedback, and we are looking forward to hearing from you.

Yours sincerely,
The authors

Reviewer comments:

Reviewer 2:

1) The paragraph beginning «Still, the characteristics...» is confusing. It starts out stating that there is variability across cultures but then discusses Norwegian IDS in particular without a comparative lens. Then appears to return to a comparative lens but with quite narrow focus.

- Thank you for bringing this to our attention. Please see page 6 for the revised paragraph. We have made this paragraph clearer by discussing how different characteristics of IDS varies across languages, mainly discussing Norwegian as compared to other languages. We have made the paragraph more nuanced by adding literature on differences in VSA across languages (e.g., hyperarticulation, hypoarticulation, and no difference between IDS and ADS).

2) In general, I found the review of paternal IDS a bit lean. For example, there are some older studies on e.g., the Father-Bridge hypothesis (see work by Tomasello, Berko Gleason) that might provide some relevant bigger-picture theoretical meat.

- Thank you for the comment and suggestions. Based on your suggestions, we have included more substantial literature on paternal IDS in the introduction. Please see pages 6 to 9 for more details.

3) With respect to the work on preference for IDS, work by Newman may be worth including.

- Thank you for the suggestion. We have included references to Newman's work in the revised manuscript, see page 12.

4) The claim on page 7 that "it is unknown whether fathers modulate their IDS" is too strong given that there is indeed existing literature on this topic.
- Thank you for the comment. This claim has been removed in the revised manuscript.

5) I believe there IS some literature on the impact of caregiver experience that could be explicitly mentioned (…)
- Thank you for bringing this to our attention. We have found a study by Weirich and Simpson (2019) and included this in the revised manuscript, see page 10.

6) Page 8: "The infants… will only differ in…" this is again too strong of a claim.
- Thank you for bringing this to our attention. The claim has been removed from the revised manuscript.

7) (…) What will the timeframe be (and how might it vary across fathers) between when the father completes the online questionnaire and when they come to the lab? Can the authors clarify "main caregiver" (both for review purposes and to ensure that the question is interpreted consistently by the fathers)?
- Thank you for the questions. We have now added more information, see pages 18-19 (time line) and page 13 (main caregiver), which will hopefully answer both questions.

8) Perhaps a copy of the actual questionnaire would be helpful?
- Thank you for the comment. The questionnaire (in English and Norwegian) is available at the OSF in the folder "Materials", link to the OSF page is https://osf.io/5qjuk/?view_only=af30057f71474783a6d7629b985fa4b1

9) Perhaps the fathers who are more comfortable around their infants are more likely to take longer leaves AND produce stronger IDS?
- Thank you for bringing this to our attention. Since being comfortable around children may come from previous experience with kids, we have decided to add a question in the questionnaire and ask whether fathers have any previous working experience with kids (e.g., as a teacher in kindergarten and in school). If they have any previous experience, they are excluded from the study. Also, we have adopted the questionnaire to include questions about paternal attitudes and knowledge on language development. This would also address your concern, as fathers who think speaking in IDS is important would be more likely to speak IDS and perhaps be more comfortable to speak it to their infants. Both reading as an activity and paternal attitudes will be explored in an exploratory analysis.

10) What happens if an infant fails to calibrate?
- Thank you for the question. We see that we have been unclear on the consequences of an unsuccessful calibration. We have added this as an exclusion criterion (see page 24), meaning that infants with an unsuccessful or incomplete calibration will be excluded from the study.

11) The decision to use word lists rather than utterances is unusual (and differs from the ManyBabies study).
- Thank you for bringing this to our attention. This was an error, and it should of course have been utterances. It has thus been edited in the revised manuscript. Please see pages 20-21.

12) Why only 8 trials?
- Thank you for bringing this to our attention. This was a typo, and it should have been 16 trials (8 trial pairs). It has thus been edited in the revised manuscript. Please see page 20.

13) p.18 The comment at the bottom of the page could use some further fleshing out (how will they be transformed? How will "normally distributed" be assessed?) and this information might be better located where the other transformations are outline, on the following page.
- Thank you for the comment. As the log transformation and deviance from normality was explained on the following page, we have removed this comment from the bottom of p.18.

Reviewer 1:
1) Perhaps the researchers could add a questionnaire to test paternal knowledge, attitudes, and/or beliefs to distinguish between these two interpretations?
- Thank you for your comment. We have edited the questionnaire to include questions about belief, knowledge and attitudes on language development in order to distinguish between paternal experience and paternal knowledge/attitudes. Please see page 19 for further description of the added questions. This measure will be added to the exploratory analyses.

2) I am a bit concerned about the choice of READING as an activity to elicit IDS. (…) Do the authors have a way of controlling how much READING fathers do with their babies and/or controlling for things like reading disorders?
- Thank you for your comment and your concern. Reading as an activity was chosen based on experience from previous studies, showing that most Norwegian parents do read to their infants and young children, and that reading does elicit differences between IDS and ADS for most acoustic features (see Rosslund et al., 2022a). In the questionnaire, there is also a question on how often the father has read to their infant the last two weeks, and this information will be reported. Fathers who do not read to their infants will be excluded from the study. We will not control for reading disorders in the current study. Please see file *Questionnaire_English_revised_after_stage1.pdf* in the folder "Materials" for the English version of the questionnaire on OSF:
https://osf.io/5qjuk/?view_only=af30057f71474783a6d7629b985fa4b1

Reviewer 3:
1) In reading the introduction several times I worried about birth order effects, but was reassured when the methods specified that all children would be first-born. It might be useful to mention this design detail earlier in the paper.
- Thank you for the comment. We have included this information earlier in the paper, please see page 5 and page 13.

2) However, one issue that has not been considered here is the measurement reliability of the infant task – how stable are individual differences as measured by the IDS-ADS preference task?
- Thank you so much for the comment. We have taken several steps to increase measurement reliability in line with your suggested research paper. First, we will compute and report the Intraclass Correlation Coefficient. Second, we will weight the contribution of the number of completed trials per child and register data points to a full-null model comparison (see Planned statistical analyses for further description). This approach will allow us to maintain a large sample size while weighting infants who contribute to more reliable data (more trials). As such, the weighted regression will take into consideration the limited reliability when the

infants don’t complete the full experiment. Third, we will only include infants who completed at least half of the trials (4 in each register). And fourth, we will clarify that when drawing interference, we will not interpret non-significant correlations, meaning that absence of correlation will not be interpreted in the current study. Please see page 17 for more information about measurement reliability of the current study.

**Fathers learning on the job: Role of Paternity Leave Duration on Paternal Infant-Directed Speech and Preference for Male Infant-Directed Speech in Infants**

Word count: 9154

## Abstract

The acoustic properties of infant-directed speech (IDS) and the functions that IDS may serve in language development have drawn noticeable interest in infant development research. However, previous research has mostly explored IDS in mothers and the preference for maternal IDS, with few studies assessing the role of exposure to or parenthood experience with an infant on acoustic properties of IDS and infants' preference for IDS. The current study will thus explore infant-directed speech in Norwegian fathers and the role of experience (duration of parental leave) on paternal language and infants' preference for male IDS. By using eye tracking technology, this study will be one of few to explore whether first-born infants prefer male infant-directed speech in early language development and if their preference is modulated by the amount of exposure to male IDS. The results of the current study will provide insights into the mechanisms affecting infant-directed speech and infants' preference for IDS in infancy.



*Keywords:* Infant-directed speech; language development; psycholinguistics; eye tracking; paternity leave

## Introduction

Infant-directed speech (IDS) is the speech produced by caregivers while interacting with their infants. IDS, as compared to adult-directed speech (ADS), is characterized by a higher pitch range (Christia, 2013), exaggerated intonation contours (Fernald et al., 1989; Fernald & Simon, 1984), and an expansion of the vowel space (Kuhl et al., 1997; Liu et al., 2003; Kalashnikova & Burnham, 2018). IDS has shorter utterances, a higher fundamental frequency, a simpler syntax and a higher repetition of single words as compared to ADS (Outters et al., 2020; Grieser & Kuhl, 1988).

The role of IDS in language development has been highly discussed. For instance, research in 7- to 8-month- old American infants has shown that infants were able to recognize words 24 hours after word familiarization when the words were produced in IDS, but not when the words were produced in ADS (Singh, Nestor, Parikh, & Yull, 2009). Thiessen, Hill and Saffran (2005) have also shown that 6- to -8-month- old infants were able to segment words from sentences only when the sentences were produced with intonation contours characteristic of IDS. Similarly, British-speaking infants were only able to segment words at 10,5 months when the sentences were produced in exaggerated IDS (Floccia et al., 2016), and word segmentation was only successful among German-speaking infants when the stimuli had acoustic properties that matched an American-English IDS (Schreiner & Mani, 2017), which is more exaggerated than German IDS. Furthermore, words produced with a wider frequency range and a higher fundamental frequency – characteristic of IDS – have been suggested to facilitate word learning in early language acquisition when infants' vocabulary sizes are relatively small (Ma, Golinkoff, Houston & Hirsh-Pasek, 2011). Other research has shown that infants who were exposed to more IDS early in life had larger expressive vocabularies at 24 months (Weisleder & Fernald, 2013), and a study by Raneri et al. (2020) found that a slower articulation rate (number of syllables per second) in IDS addressed to infants when they were 7 months significantly correlated with later vocabulary size when the infants were two years. As such, extensive research suggests that IDS plays an important role in language development.

Still, the characteristics of IDS vary across languages and cultures. For example, in Norwegian, in relation to the current study, some research indicates that IDS has longer vowel duration as compared to ADS, although the difference in vowel duration between IDS and ADS decreases during the first six months of the infants' life (Englund & Behne, 2006). Norwegian IDS has also shown to have higher pitch and greater variation in vowel pitch as compared to ADS (Steen & Englund, 2022). This is in support of similar findings found in other languages (Cristia, 2013; Fernald & Simon, 1984; Fernald et al., 1989; Marklund & Gustavsson, 2020). On the other hand, Steen and Englund (2022) found that Norwegian pedagogical employees had smaller vowel space areas in IDS than in ADS and that vowels were in fact hypoarticulated in IDS as compared to ADS. Another example comes from Dutch: Benders (2013) found that Dutch mothers had smaller vowel space areas when talking to their infants. These findings do not support previous findings of vowel hyperarticulation in IDS reported in other languages (Christia & Seidl, 2014; Kuhl et al., 1997; Liu et al., 2003). A study by Rosslund et al. (2022a), however, found an expanded vowel space area in both maternal and paternal Norwegian IDS, as compared to ADS, as well as more variable vowel categories, higher pitch, wider pitch range and longer vowel duration. Differences in infants' age or dialectal variation, however, may account for these differences in vowel space area, as most participants in the studies by Englund and Behne (2006) and Steen and Englund (2022) spoke a Central Norwegian dialect, while the participants in the study by Rosslund et al. (2022a) spoke an Eastern Norwegian dialect. Moreover, other studies have found no differences in vowel space area between IDS and ADS in American-English (Burnham et al., 2015) and Cantonese (Xu Rattanasone et al., 2013). This suggests that there are language- and culture-specific, and perhaps situational, factors affecting speech modulation in infant-directed speech.

Although research shows that the characteristics of IDS may vary across languages, this research is almost exclusively based on maternal IDS. It is thus not fully known whether previous findings on Norwegian IDS, and IDS in general, is applicable to paternal IDS. Yet, in many countries (including Norway), fathers play an important role of infants' upbringing and spend up to 6 months with the infant while on paternity leave during the infant's first year. To fully capture the language environment of the modern-day infant and understand language development, it is thus necessary to include fathers in research studies (Ferjan Ramírez, 2022).

A reason to why research studies have focused on maternal speech may be due to larger amounts of IDS infants and young children hear from their mother as compared to their father. Bergelson et al. (2019) found that the North American children heard 2-3 times more IDS from females than from males, and that children heard increasingly more IDS as they grew older. Shapiro et al. (2021) found similar results, with infants in English-speaking families being exposed to 46.8 % less words and 51.9 % less IDS from fathers than from mothers. Furthermore, they found that both paternal and maternal IDS increased from 6 months to 24 months, but the rate of increase was 2.8 times faster in fathers as compared to mothers. This suggests that the quantitative gap in IDS between mothers and fathers may be larger in early infancy, perhaps as a result of fathers spending more time interacting with their children in more physical activities later in the infants' development (Shapiro et al., 2021; Ferjan Ramírez, 2022).

Still, there is some research on paternal IDS. Research studies investigating the quality of paternal IDS have found mostly similarities between IDS in mothers and fathers (Hladik & Edwards, 1984; Golinkoff & Ames, 1979; Fernald et al., 1989; Weirich & Simpson, 2019; Rosslund et al., 2022a; Jacobson et al., 1983; Papoušek et al., 1987). For example, Fernald et al. (1989) found that both mothers and fathers had a higher mean pitch, greater pitch variability, shorter utterances and longer pauses in IDS as compared to ADS, but only mothers had a wider pitch range when talking to preverbal infants. These findings were consistent across languages, including French, German, Italian, British and American English, and Japanese (Fernald et al., 1989). Benders et al. (2021) recently found similar results in Dutch, showing that both mothers and fathers raised their average pitch, expanded their pitch variability within utterances and increased their pitch variability across utterances in IDS. Fathers, however, increased their pitch variability both across and within utterances more than mothers, suggesting that paternal IDS may be more dynamic and energetic as compared to maternal IDS (Benders et al., 2021). Furthermore, Gergely et al. (2017) found that Hungarian fathers' speech was more sensitive to the infant's age, as compared to mothers' speech, where fathers used significantly higher pitch and a broader pitch range when speaking to younger infants than to older infants and toddlers. They also found that both parents hyperarticulate their vowels when addressing their infant, mothers more than fathers for infants under 18 months of age, but not when addressing their pet dog, suggesting that hyperarticulation may be related to language tutoring and language development. On the other hand, Rosslund et al. (2022a) found that Norwegian mothers, but not fathers, had longer vowel durations in IDS as compared to ADS, all of these findings suggesting that there are cross-gender differences in acoustic measures in speech addressed to infants.

Although studies have found similarities between maternal IDS and paternal IDS, a study by Bingham et al. (2013) found differences in language use between mothers and fathers across contexts and settings. In their study, mothers' language use in a triadic mother-father-infant setting was predicted by maternal education, child's age and maternal employment status, while fathers' language use in the same setting was related to the child's age, balanced co-parenting and paternal sensitivity. Similar results were found in an earlier study by Golinkoff & Ames (1979), where fathers took less conversational turns and spoke less in triadic settings as compared to a dyadic setting (father-infant). These findings suggest that fathers may feel less responsible for the interaction if the mother is present or that it may be easier for the mother to interact with the infant as a result of more experience and time with the child. Furthermore, research shows that mothers often talk more to their infants and young children, and that mothers often talk more to daughters than to sons (Leaper et al., 1998). Some research also suggest that fathers may demand more of their children conversationally by producing more wh-questions, more imperatives and more frequent requests (Rowe et al., 2004; Leaper et al., 1998; Gleason, 1975). In her study, Gleason (1975) discussed how a father may serve as a bridge to the adult world by providing his child with more experience with demanding conversations, leading to the controversial Father-bridge hypothesis. Still, the family roles at the time of Gleason's study were very different from the modern-day Norwegian families, and Gleason found that family roles indeed were reflected in the fathers' language: "Finally, the fathers' language clearly demarked their role within the family: a father playing with his small son, for instance, might break off the game to send the child to his mother to have his diaper changed" (Gleason, 1975, paragraph 5). Similar results were found in a study by Le Chanu & Marcos (1994), where the differences in vocabulary and conversational aspects (e.g., the content of questions, if the parent understood their child's utterances, and if the parent followed the child's topic of interest) were explained in terms of parental roles and how mothers' role is to "provide a feeling of security" while fathers' role is to prompt the child to attain higher levels of success.

With few studies investigating the role of experience and duration of the paternity leave on fathers' speech when interacting with their child, an important question in the current study is thus whether Norwegian fathers also adapt their speech when interacting with their infants and whether this adaptation is modulated by their experience with their child as the main caregiver. In the study by Jacobson et al. (1983), they found that non-parents with little prior experience with children still modified their fundamental frequency (f0) as much as the parents in study, suggesting that certain acoustic features of IDS may be attributable to something other than experience. Still, it is not known whether the fathers in the study by Jacobson et al. (1983) were engaged in caregiving activities or had caregiver responsibilities. A newer study by Weirich and Simpson (2019) found that there was no significant effect of gender or parental involvement on German fathers' IDS, suggesting that fathers who are more involved in child care do not modify their speech significantly more than less involved fathers. Although the fathers in the latter study were more involved in child care than the control group (fathers who were less involved in child care), the distribution of involvement in child care shows that the fathers were still considerably less involved in child care than mothers, suggesting that mothers still had most of the caregiver responsibilities in the infants' first year of life. Based on previous literature on IDS, in the current study, we expect that fathers will adapt their speech similarly to mothers provided they have had enough learning experience as the main caregiver and spent enough time with their child. More specifically, we predict that more experience with the child will result in higher adaptation in IDS. In fact, recent research shows that parents fine-tune their speech to their child according to their child's development, suggesting that IDS may serve as a way of fine-tuning the complexity of the parents' speech in relation to the skills of their children (Leung, Tunkel & Yurovsky, 2021). Such parental scaffolding would require the parents to have an awareness of the skills and development of the child, which would mainly be acquired through experience with the child and experience as a caregiver. IDS, as a method of parental scaffolding, may thus explain the variation in linguistic properties across languages and throughout children's development.

Regardless of the cross-linguistic differences in the acoustic features of IDS, overall, research suggests that some of the main characteristics of IDS, such as vowel hyperarticulation, pitch, repetition of words and a simpler syntax, among others, may serve different functions at different stages and that IDS may have both attentional and linguistic functions (Kuhl et al., 1998; Liu et al., 2003; Cristia, 2013; Outters et al., 2020; Grieser and Kuhl, 1988; Kalashnikova & Burnham, 2018). For instance, a study by Kitamura and Burnham (2018) indicated that parents may use IDS to show positive affect, express affection, comfort or soothe, to encourage attention, and to direct behavior. They found that IDS with the intent of providing comfort or to soothe is more evident at birth, and that IDS with the intent to direct behavior is more prevalent when the infants are around 9 months old. Furthermore, mean pitch (f0) was mostly associated with affective-type utterances, while pitch range was mostly associated with utterances of a more directive intent (to encourage attention or to direct behavior). Benders (2013) found that Dutch mothers consistently raise the formant values F2 and F3 of the corner values and the spectral mean of the voiceless fricatives, which she argues are acoustic markers of positive affect. Benders thus hypothesizes that IDS may be a side-effect of smiling or stemming from the articulatory means the mother does in order to convey positive emotions and make her voice less threatening (Benders, 2013). In the study by Raneri et al. (2020) the authors also found that mothers' speech rate increased as their infants got older, suggesting that parents do modulate their speech in relation to their infants' age and development. In sum, IDS may have different functions at different stages of development, and these functions may be visible in the acoustic properties of IDS. Still, it is not yet known, to the best of our knowledge, whether fathers, similarly to mothers, modulate their speech when talking to their infants, and if the modulation is visible in the typical acoustic markers of IDS. Furthermore, few studies have taken into account the role of experience and whether experience as the main caregiver affects speech modulation at different stages of child's development. The current study will address these matters by assessing Norwegian IDS among fathers with varying lengths of paternity leave during their infants' first 8 months of life.

**Preference for infant-directed speech in infancy**

While the role of IDS in language development has been highly debated in recent research, there is extensive body of research suggesting that young infants prefer IDS over ADS (Cooper & Aslin, 1990; Pegg, Werker & McLeod, 1992). Cooper and Aslin (1990) found that both 1-month-old and 2-day-old infants fixated longer at a visual stimulus if the fixation produced IDS audio as opposed to ADS audio, suggesting that the preference for IDS may even be present from birth. Outters et al. (2020) suggested that the preference for IDS may vary across development, where older infants do not show a preference for IDS over ADS. Furthermore, they found that the degree of IDS preference may be related to the quality of maternal IDS that the infant had been exposed to earlier in life. Similarly, Newman et al. (2004) found that the youngest infants (4 months) in their study had a preference for IDS over ADS, while 9-month- old and 13-month- old infants did not. Moreover, none of the age groups had a greater preference for IDS when listening to IDS with a noisy background as compared to IDS in quiet, suggesting that infants in general prefer to listen to IDS in quiet settings (Newman et al., 2004). Hayashi, Tamekawa and Kiritani (2001), however, found a U-shaped pattern of preference for IDS in Japanese infants, where the youngest (4-6 months) and the oldest (10-14 months) infants showed a preference for IDS, while the infants aged 7-9 months did not show a preference. As such, the results on IDS preference are mixed and may differentiate according to the methodology and the language being tested in the experiment. Therefore, The ManyBabies Consortium (2020) assessed IDS preference using several methodologies in a large study with 2329 infants from 67 labs in North America, Europe, Asia and Australia using North American English IDS. They found that the IDS preference was in fact stronger in older infants than in younger infants, and that infants had a stronger preference for IDS if the stimuli were presented in their native language (The ManyBabies Consortium, 2020). This suggests that IDS preference increases with age, but it is unknown whether increased preference is related to infants' maturation or to their increased exposure to IDS. Hence, beyond the interest of evaluating preference for male IDS and whether fathers fine-tune their speech with experience, the current study will also address the issue of whether preference for male IDS increases with more exposure to male IDS.

## The Current Study

Most research on infant-directed speech is on mothers' speech and it is not fully known whether fathers modulate their IDS when speaking to a child and whether it is modulated by the amount of experience with the child. There is also very little research on preference for male infant-directed speech in infancy and the role of exposure to male IDS. The present study will explore whether fathers modulate their speech when talking to infants, and if this modulation is related to their experience as the main caregiver. To assess speech modulation, we will record fathers reading to their infant (IDS) and to the researcher (ADS). By using eye tracking technology, we will also explore whether first-born infants prefer male infant-directed speech over male adult-directed speech in early language development, and whether this preference is modulated by the amount of exposure to parental speech in infancy.

Parental leave in Norway is articulated into three phases; a mother-specific leave of 3+15 weeks (the first 3 weeks being before birth), followed by 16 weeks of shared leave (where parents decide on the split of the time) and finally, when the child is 31 weeks old, a father-specific leave of 15 weeks, the father-specific leave being one of the longest in the world (OECD, 2021). Consequently, the total duration of father-specific leave ranges between 15 and 31 weeks, and starts as early as when the child is 15 weeks or as late as 31 weeks depending on parental distribution of the shared leave (NAV, 2022b). As such, it offers an opportunity to gain knowledge on the role of father-specific leave duration on paternal infant-directed speech. In the current study, we will examine paternal speech and infants' preference for male infant-directed speech in 70 Norwegian monolingual father-infant dyads (see Methods for sample size rationale). The infants will thus have the same age (8 months +/- two weeks), but will differ in the amount of time their father has spent with them as the main caregiver. Here, main caregiver denotes the person that is mostly at home with the infant having caregiver responsibilities. In the case of the current study, fathers in paternity leave with 100 % coverage will be defined as the main caregiver, although it is acknowledged that the mother of the infant naturally will have caregiver responsibilities regardless of the parental leave status. For example, the mother will likely be present in the evenings and outside of working hours.

Speech modulation will be assessed by acoustically analyzing speech recordings in IDS and ADS using the Praat Software (Boersma & Weenink 2022). Preference for male infant-directed speech will be tested using an Eyelink 1000 Plus eye tracker to measure the infants' looking times at the checkerboard screen while listening to male speech in either IDS or ADS. Significantly longer looking times in IDS conditions compared with the ADS conditions will be interpreted as evidence for infants' preference for male IDS.

Acknowledging the limited measurement reliability of the infant task, we will apply a conservative approach while interpreting the results and will only interpret correlations differing significantly from zero. Absence of a correlation will thus not be interpreted as lack of a relationship between the amount of experience with the child and the degree of modulations in IDS vs ADS in the current study.

## Hypotheses

The following hypotheses are considered in the current study (see also TABLE 1 for study design table):

### *Paternal infant-directed speech*

H1a. If fathers modulate their speech when talking to infants, then fathers' IDS will be more pronounced (larger vowel space area, higher f0 mean, wider f0 range, slower articulation rate) as compared to fathers' ADS.

H1b. If paternal speech modulation is related to experience as the main caregiver, then paternal IDS will be more pronounced when fathers have had a higher number of days since the beginning of their paternity leave.

### *Preference for male infant-directed speech in infancy*

H2a. If infants prefer male IDS over male ADS, then, in the eye tracking experiment, infants will have a longer looking time in IDS trials than in ADS trials.

H2b. If IDS preference is modulated by the amount of exposure to parental speech, then infants will have a higher preference for male IDS when their father has had a higher number of days since the beginning of their paternity leave (increased exposure to male language input).

TABLE 1

| **Question** | **Hypothesis** | **Sampling Plan** | **Analysis Plan** | **Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis** | **Interpretation given different outcomes** | **Theory that could be shown wrong by the outcomes** |
|---|---|---|---|---|---|---|
| Do fathers modulate their | H1a. If fathers modulate their | Power analysis | Full-null model comparison. | Effect size was obtained from | If the full-null model | Support for H1a will suggest that |

| | | | | | | |
|---|---|---|---|---|---|---|
| speech when talking to infants (larger vowel space area, higher f0 mean, wider f0 range, slower articulation rate)? | speech when talking to infants, then fathers' IDS will be more pronounced (larger vowel space area, higher f0 mean, wider f0 range, slower articulation rate) as compared to fathers' ADS. | | Null model: Acoustic measure ~ SES + (1+Register\|Participant)<br><br>Full model: Acoustic measure ~ Register*Pat_duration + SES + (1+Register\|Participant) | the study by Rosslund et al. (2022b) for each acoustic measure. We computed the power analysis in G*Power (Faul et al. 2009) to find the minimum sample size based for the design to be sensitive enough to detect all effect sizes. | comparisons for all acoustic measures are non-significant, it will disconfirm the hypothesis. If the full-null model comparison is significant, we will interpret the predictors in the model.<br>A significant effect of register on acoustic measure(s) will confirm the hypothesis. | fathers' IDS differs from ADS (aligning with the literature on mothers' IDS), while evidence against H1a would suggest fathers' IDS does not differ from ADS. In the presence of a significant interaction, the interpretation of the main effect of Register would be limited. |
| Is paternal speech modulation related to experience as the main caregiver? | H1b. If paternal speech modulation is related to experience as the main caregiver, then paternal IDS will be more pronounced when fathers have had a higher number of days since the beginning of their paternity leave. | Power analysis | | Same as for H2b: We computed the achieved power using G*Power (Faul et al., 2009) with a sample size of 70 for the most complex model (a conservative approach) with two main effects and an interaction term. | A significant effect of paternity leave and/or its interaction with register will provide evidence that paternity leave modulates the acoustic measure and will confirm the hypothesis. No significant effect will disconfirm the hypothesis. | Support for H1b would suggest that fathers' accumulated experience as the primary caregiver is associated with the difference across registers (the adaption of IDS to the child), while evidence against H1b would fail to support that claim. |
| Do infants prefer male IDS over male ADS? | H2a. If infants prefer male IDS over male ADS, then, in the eye tracking experiment, infants will have a longer looking time in IDS trials than in ADS trials. | Power analysis | Full-null model comparison<br><br>Null model: Looking time ~ SES + (1\|subject)<br><br>Full model: Looking time ~ Trial_Type*Pat_Duration + SES + (1\|subject) | Effect size estimated were obtained from the study by The ManyBabies Consortium (2020), and G*Power was used to calculate sample size. | If the full-null model comparison is significant, we will inspect the predictors (trial type, duration of paternity leave and their interaction) to assess which ones are driving the effect. A significant effect of trial type in the absence of the interaction will confirm the hypothesis. No significance of trial type will not be interpreted. No significant effect will disconfirm the hypothesis. | Evidence for H2a would suggest that 8-month-old infants attend longer to male IDS than ADS (as for female IDS). In the presence of a significant interaction, the interpretation of the main effect of Trial Type would be limited. |
| Is (male) IDS preference modulated by the amount of exposure to parental speech? | H2b. If IDS preference is modulated by the amount of exposure to parental speech, then infants will | Power analysis | | Same as for H1b: We computed the achieved power using G*Power (Faul et al., 2009) with a | If the full-null model comparison is significant, we will inspect the predictors to see which ones are | Evidence for H2b would suggest that infants' experience with and exposure to a male primary |

| | | | | | | |
|---|---|---|---|---|---|---|
| | have a higher preference for male IDS when their father has had a higher number of days since the beginning of their paternity leave (increased exposure to male language input). | | | sample size of 70 for the most complex model (a conservative approach) with two main effects and an interaction term. | driving the effect. A significant effect of paternity leave duration or the interaction between trial type and paternity leave duration will confirm the hypothesis. No significance of paternity leave duration or the interaction will not be interpreted. | caregiver is associated with their preference for male IDS. If there is a positive main effect of paternity leave duration, it would suggest that infants attend to a male voice longer, regardless of register, when the father has spent more time as the main caregiver. A negative main effect of paternity leave duration would suggest that longer paternity leave coincides with a reduced preference for IDS. Evidence against H2b will not be interpreted. |

## Methods

### Participants

To determine the maximum sample size for the current study, we first obtained, for the hypothesis H1a, the effect sizes reported in previous studies. For the H1a, the effect sizes for the acoustic measures associated with the differences between the IDS and ADS in Norwegian fathers were: mean pitch with *g*Hedges = −0.85 ((95% CI = −1.36 to −0.36), pitch range with *g*Hedges = −0.47 ((95% CI −0.93 to 0.03), full vowel space with *g*Hedges = −0.53 ((95% CI −0.99 to −0.08), and articulation rate with *g*Hedges = 0.51 ((95% CI 0.06 to 0.96) (Rosslund et al., 2022b). Using G*Power (Faul et al., 2009) to compute sample size with a power of 80 % and a significance level of 0.05 for H1a, the current study will need 13, 38, 30, and 33 fathers to detect effect sizes of -0.85, -0.47, -0.53 and 0.51, respectively. Based on H1a only, the sample size would then be 38 participants to detect all effect sizes with a power of 80 %.

To compute the maximum sample size to test H2a, effect size estimates were obtained from the study by The ManyBabies Consortium (2020). Here, the mean effect-size estimate for infants' preference for maternal IDS vs. ADS was Cohen's d = 0.35 ((95% CI = 0.29 to 0.42), z = 10.67, p < .001). Using G*Power (Faul et al., 2009) to compute sample size with a power of 0.80 and a significance level of 0.05 for H2a, the current study will need a sample size of 67 participants (infants) to detect an effect size of 0.35. Based on these computations, the maximum sample size of the current study will be 70 participants (father-infant dyads) resulting in 140 recordings (70 in each register). For the hypotheses H1b and H2b, we will run mixed-effects regression models. Since there is no known effect size of the interaction between the IDS/ADS differences and duration of paternity leave (the main hypothesis for H1b) or the IDS preference and paternity leave duration (the main hypothesis for H2b), we computed the power that would be achieved with a sample size of 70 and for the most complex model (so we applied a conservative approach) that would contain two main effects and the interaction term, i.e., for the H2b. A computation of achieved power in G*Power (Faul et al., 2009) for the most complex model with the statistical test "linear multiple regression $R^2$ increase" using the least meaningful effect size of $f^2$=0.30, 1 as the number of tested predictors, 3 as the number of total predictors and a sample size of 70 showed an achieved power of 89 %, suggesting that the design is powerful enough to detect the interaction effect.

70 father-infant dyads will thus be recruited to participate in the study. The participants will be recruited from the National Population Registry (Folkeregisteret), and invitations will be sent by postal services to all families living in the Oslo area with infants approaching 8 months of age. In the invitation letter, they will be informed of the inclusion criteria (see below) and asked to sign up for the study if they want to participate. If father-infant dyads are later excluded from the study (see exclusion criteria), we will recruit additional participants to reach the total of 70 father-infant dyads.

### Measurement reliability

Acknowledging the limited measurement reliability of the infant task, we have adopted several solutions from Byers-Heinlein et al. (2022) to increase the measurement reliability in the current study. First, we will compute and report the Intraclass Correlation Coefficient using the function ICC3k (a multiple measures variant of a 2-way random-effects model) in the *psych* package of the R software. Second, to account for the variation in the number of observations per individual and register (IDS vs. ADS) in the infant preference task, we will weight the contribution of the data points to the model by the number of completed trials per child and register, and consequently, the more trials an infant has completed, the higher its contribution to the model (see Planned statistical analyses).

### Inclusion criteria

The following criteria will be used to include fathers and infants: (1) the father has started his paternity leave at the time of data collection; (2) the father has not been the main caregiver the first 5 months (except the first two weeks after birth when both parents may stay at home); (3) the father must use the father-specific weeks of the parental leave at one go and have no part-time leave; (4) the mother and father have lived together up since the birth of the infant and until the time of the data collection (5) the child was born full term (gestational weeks >37); (6) the child is exposed to 90% Norwegian or more at home; (7) both parents speak Norwegian to the child; (8) the child has no known hearing or visual impairments; and (9) it is their firstborn child. The study has been approved by the Norwegian Centre for Research Data (NSD), and has been recommended by the Internal Ethics Committee at the Department of Psychology at the University of Oslo.

## Stimuli

### Paternity leave and language background questionnaire

Prior to the visit to the lab, the fathers will fill in a questionnaire regarding paternity leave and language background. All fathers will receive a link to the questionnaire one week before the lab meeting, or less than a week if the lab meeting is scheduled sooner. The fathers may fill in the questionnaire up until the lab meeting, meaning that all questionnaires have been filled out between one week prior to the lab visit and up until the time of the scheduled meeting. This questionnaire will ask fathers to provide information about their paternity leave by asking two questions: “When did your paternity leave start?” and “Have you had a longer period of time (more than 7 days) before your paternity leave where you were the main caregiver for your child?”. Respectively, the father will provide the date of the start of their paternity leave and answer yes/no.  The first question will provide information that will be used as the independent variable (number of days in paternity leave will be calculated using the start date of paternity leave until the date of the data collection). Based on the data from the Norwegian Welfare and Labour Administration (NAV, 2022b), it is expected that most fathers in Oslo will only have the father-specific weeks or the father-specific weeks plus a minor fraction of the shared leave. An inclusion criterion is thus that the father must, as a minimum, use the father-specific weeks to participate in the study. Furthermore, it is possible for Norwegian fathers to postpone their paternity leave up until their child is three years of age, and/or split the paternity leave to several time periods, and combine the paternity leave with work. It is thus required in study that the father has *not* postponed his paternity leave, that the paternity leave has not been interrupted/split up until the time of data collection, and that the father has not been working at the same time as his paternity leave. If all fathers only have the father-specific weeks (and none of the shared period of the total parental leave) after the first wave of data collection, then the second wave of data collection will target fathers who have had a minimum of two weeks of the shared period.

The questionnaire will also collect information regarding the language environment of the infant (the parents' language(s), the infants' language(s), and the parents' educational level (that will be used as a control variable). Furthermore, the fathers will be asked whether they have had any previous working experience with kids, for example as a teacher in kindergarten or in school. If they answer yes to this question, we will exclude them from the study. Also, the questionnaire will collect information about the fathers' attitudes and beliefs on language development and language learning. These questions include (translated): "Parents may learn babies to talk by talking with them", "Reading to a child is of no use as long as the child has not learned to speak yet", "It is important to not talk baby talk when talking to a little child", "I automatically use baby talk (e.g. words like "pipp-pipp" and "vov-vov") when I talk with a little child", "When I speak with a little child, I often use a different voice with a more lively tone", and "When I speak with a little child I often speak slower and clearer". The fathers will be able to answer these questions on a Likert-scale from 0 (do not agree) to 6 (very much agree), and a sum score (note that the score of question 2 and 3 will be reverted) will make up the paternal attitudes measure in the exploratory data analysis.

The fathers in the study will also be asked to provide information on how often they read with their child in the past two weeks. If they did not read anything at all, they will be excluded from the study (please see the OSF for the questionnaire:

https://osf.io/5qjuk/?view_only=af30057f71474783a6d7629b985fa4b1, file name *English_questionnaire_revised_after_stage1.pdf* in the folder "Materials"). The fathers' response to the question on reading will make up the reading activity measure in the exploratory analysis.

**Recordings of IDS and ADS**

The two registers, IDS and ADS, will be assessed by recording the fathers' voice when reading twice (to their child and to an adult) a short story from a custom-based children-friendly book. The same custom-based book as in the study by Rosslund et al. (2022a) will be used, containing all 9 Norwegian long vowels (/iː/ /yː/ /eː/ /øː/ /æː/ /ʉː/ /uː/ /ɔː/ and /ɑː/) presented in 5 different words and repeated 2 times. All vowels will thus be encountered 10 times each during five short stories (SEE TABLE 2). These five short stories are presented on five pages with colorful illustrations, and contains in total 39 sentences, 327 words and 90 target words with target vowels (SEE TABLE 3). The target words are monosyllabic and bisyllabic lexical and function words, and each word is repeated twice.

TABLE 2

| /iː/ | /yː/ | /eː/ | /øː/ (eu) | /æː/ (ae) | /ʉː/ (uu) | /uː/ | /ɔː/ (o) | /ɑː/ |
|---|---|---|---|---|---|---|---|---|
| bil | lys | se | brød | der | lue | bok | sove | banan |
| gris | fly | skje | snø | her | pute | sko | tog | bade |
| spis e | dyne | mer | dør | være | ku | fot | hår | kake |
| skiv e | dyr | nese | bjørn | bære | mus | sol | måne | mage |
| vi | ny | lese | løpe | skjær e | fugl | hallo | gå | bra |

TABLE 3

| **Original** | **English translation** | **Phonetic transcription** |
|---|---|---|
| Der ute skinner **solen** og **fuglene** kvitrer. Det er ganske kaldt og **bjørnen** har tatt på seg **lue** og **sko**. Det er deilig å være ute når det er **snø** på bakken. **Bjørnen** børster bort **snøen** fra **nesen** og den **hårete** pelsen på **magen**. Men det kommer bare enda **mer**. Han må børste **nesen** og den **hårete** pelsen en gang til. | Out there, the **sun** is shining and the **birds** are tweeting. It is quite cold and the **bear** has put on a **hat** and **shoes**. It is nice to be outside when there is **snow** on the ground. The **bear** brushes the snow away from his **nose** and the **hairy** fur on his stomach. But it keeps coming even **more**. He has to brush his **nose** and the **hairy** fur once more. | Der ute skinner **/su:lən/** og **/fʉ:ləne/** kvitrer. Det er ganske kaldt og **/bjø:ɳən/** har tatt på seg **/lʉ:e/** og **/sku:/**. Det er deilig å være ute når det er **/snø:/** på bakken. **/bjø:ɳən/** børster bort **/snø:ən/** fra **/ne:sən/** og den **/hɔ:rəte/** pelsen på **/mɑ:gən/.** Men det kommer bare enda **/me:r/**. Han må børste **/ne:sən/** og den **/hɔ:rəte/** pelsen en gang til. |

### Central-Fixation Eye Tracking Procedure

In order to assess preference for male infant-directed speech, the same methodology as in the ManyBabies1 project (The ManyBabies Consortium, 2020) will be adopted. In the experiment, a central-fixation eye tracking procedure using Eyelink 1000 Plus will test whether infants express more interest (here: look at a screen in front of them and not look away) when listening to audio of male IDS as compared to listening to audio of male ADS. There will be mixed trials to exclude a potential effect of the order of presentation. To create the IDS and ADS stimuli for the current eye-tracking procedure, we will follow closely the same procedure for stimuli creation as in the ManyBabies1 study (The ManyBabies Consortium, 2020). We will then create utterances in both registers, including all target words repeated twice. This will result in 16 trials (8 in each condition), lasting for 18 seconds each. The total time for all trials is 288 seconds. Also, similarly to the ManyBabies1 project (The ManyBabies Consortium, 2020), the experiment will contain two warm-up trials lasting for

18 seconds with piano music as the auditory stimulus and the same visual stimulus as the test trials.

Three external raters will rate the utterances on whether they believe the utterances are directed at infants or at adults, in order to make sure that the IDS and ADS utterances differ enough in their acoustic properties and are perceived as either IDS or ADS.

A visual stimulus in the form of a colorful checkerboard will be showing on the screen when the utterances are presented. Before each trial, a colorful small spinning circle in the middle of the screen will be displayed on a black background along with a short sound in order to gain or regain the infant's attention during the experiment.

**Procedure**

The data collection will be performed in a single session in the Babyling laboratory at the Department of Psychology at the University of Oslo. Prior to the visit, the fathers will have received an information letter with information about the study by email. They will also have received a participant number and a link to the paternity leave and language background questionnaire. The questionnaire is an online form provided by the University of Oslo: https://nettskjema.uio.no. In this questionnaire, the fathers will be asked to fill in their participant number and provide informed consent to participate in the study. The participant number will enable us to connect the information collected in the questionnaire with the information collected in the laboratory.

The researcher, a female native speaker of Norwegian, will welcome the father and infant in the reception area of the lab. Here, they will receive brief information about the study, as well as information about the following eye tracking session and recording sessions. The researcher will then lead the father and infant to the eye tracking session in the room next door.

The eye tracking session will be performed using an Eyelink 1000 Plus and an arm mount to easily position the screen and eye tracker in front of the infant. The infant will be seated in a car seat facing approximately 60 cm from a screen with 1920 x 1080 pixels screen resolution. The father will be seated directly behind the infant at all times and will be wearing headphones with masking music. He will also be asked not to point to the screen or talk to his child during the experiment. The researcher will be seated in the same room behind the infant, outside of the infant's view.

The infant will wear a small sticker on his/her forehead for the eye tracker to track the eye correctly. Before the experiment, a calibration and validation procedure will be performed where the infant will look at small blinking targets on a black background positioned sequentially on the sides of the screen (left, right, top, bottom). The validation procedure will look identical as the calibration procedure to the infant and will confirm that the calibration of the eye tracker successfully captured the eye and calculated the eye gaze accurately. The calibration and validation will be kept brief.

After the calibration and validation, the central-fixation eye tracking experiment to test male IDS/ADS preference will start. The same procedure as in ManyBabies study (The ManyBabies Consortium, 2020) will be adopted for the visual stimuli. Before each trial, a colorful small spinning circle in the middle of the screen will be displayed on a black background along with a short sound in order to gain or regain the infant's attention during the experiment. When the infant orient his/her gaze towards the screen, a visual stimulus in the form of an image of a colorful checkerboard will appear. Simultaneously, an auditory stimulus (utterances) will be played through two speakers positioned at the left and right sides of the screen. Each utterance contains 8 words and will be spoken by male voices in either IDS or ADS with an average amplitude of 65 dB and will be played until the maximum trial length of 18 seconds or until the infant has looked away for more than 2 seconds. If the maximum trial length is reached or the infant has looked away for more than 2 seconds, the attention getter will be displayed until the infant fixates back on the screen. Then the next trial will start. The experiment ends when all 16 trials have been presented to the infant.

After the experiment in the lab, the father and infant will be followed back to the reception area where they are able to debrief and ask questions before the recording sessions. The same procedure as in the study by Rosslund et al. (2022a) will be applied. Both recording sessions will take place in the reception area, and a zoom handy recorder model h4n, serial no. 00251740 will be used in both sessions. During the IDS session, the father will be instructed to read the child-friendly short-story book to his infant as he would naturally do at home. During the ADS recording, the father will read the same short-story book to the researcher. The order of recordings will be counterbalanced between participants to ensure that familiarization with the book does not impact speech performance.

After the eye tracking session and the recording sessions are finished, the infants may choose a toy of their liking from a selection of toys as a token of appreciation. The infants will also receive a diploma.

**Data preprocessing**

The recordings in IDS and ADS will be acoustically segmented and analyzed using the Praat Software (Boersma & Weenink, 2022). First, the researcher, a native Norwegian speaker, will segment all target vowels. As all participants will be male, the formant values will be extracted below a ceiling value of 5000 Hz.

Vowel segmentation will be based on these criteria: (1) vowel onset point (VOP) is at the first upward crossing in the speech signal after the release of the preceding consonant (Cristia & Seidl, 2014); (2) vowel end point (VEP) is the first downward crossing (attenuation of energy) in the speech signal after VOP *and/or* where the formant tracks for F2 and F3 is no longer visible in the spectrogram. All vowels will be included in the analysis independent of vowel duration. Exclusion criteria for vowels are as follows: (1) The vowel is interrupted by background noise, interference or talker overlap; (2) the target word has been whispered or heavily glottalized (Cristia & Seidl, 2013) or the speaker has a creaky voice or there is a heavy puff of air during the vowel (Englund & Behne, 2005); (3) the formants are not clearly visible in the spectrogram; or (4) when it is not possible to determine the onset and/or offset of the vowel.

A Praat script (Lennes, 2017) will be used to identify and compute formant values (F1-F3) for all target vowels at the mid point of each segmented vowel. In addition, another script (Hirst, 2022) will extract f0 and duration. As in Kalashnikova and Burnham (2018) and Rosslund et al. (2022a), we will convert all Hz values to semitones for f0. Articulation rate will be assessed by using a script from Rosslund et al., (2022b) to extract the number of syllables per second in each phrase.

In contrast to previous research, each vowel will also be inspected and manually corrected if the formant values were incorrectly extracted. The manual correction will be based on these criteria: (1) the mid point of the segmented vowel has one or more improbable red speckles in the formant contour of the spectrogram (Boersma & Weenink, 2022) that is clearly inconsistent with the rest of the vowel; and (2) the red speckles in the mid point of the formant contour is not making up a stable portion of the segmented vowel. In these cases, a manual correction will be performed by extracting the formant values from a stable portion closest to the mid point of the segmented vowel. If this is not possible, the segmented vowel will be excluded.

**Removing outliers (formant values)**

Formant values will be excluded from the data set if the formant values for the particular vowel is improbable. The exclusion will be based on a set of criteria (see the OSF for a full description of exclusion criteria of vowel tokens: https://osf.io/5qjuk/?view_only=af30057f71474783a6d7629b985fa4b1, file name *Exclusion_criteria_vowel_formants.pdf* in the folder "Materials" ).

**Exclusion criteria**

Father-infant dyads may be excluded from the study according to any of the following exclusion criteria: (1) less than 4 trials in each condition in the central-fixation eye tracking experiment was completed; (2) The recordings in either IDS or ADS include less than four (of five) short stories or are not recorded in entirety; (3) The recordings in either IDS or ADS are missing all formant values for any one target vowel; (4) the father did not complete the questionnaire prior to the visit to the lab; (5) the calibration of the eye tracker was incomplete or unsuccessful (3 of 5 calibration dots were not completed); (6) the father has had previous working experience with kids, for example as a teacher in kindergarten or school; (7) the father has had more than two weeks of paternity leave before his current paternity leave period (excluding two weeks birth leave); or (8) the father has reported that he did not read to his infant at all the last two weeks.

**Dependent measures**

*IDS/ADS recordings*

Formant values will be obtained from IDS and ADS recordings using a script (Rosslund et al., 2022b) based on the maximum ceiling approach mentioned in the study by Chládková, Escudero & Boersma (2011). Previously mentioned scripts (see data preprocessing) will assess f0 and articulation rate. Vowel space area, mean f0, f0 range, and articulation rate will make up the acoustic measures in IDS and ADS that will be used in the data analysis.

*Central Fixation Eye Tracking procedure – looking time at screen*

Our outcome measure for H2a and H2b is the looking time (LT) collected from the central fixation eye tracking experiment. LT is defined as the number of milliseconds when the child is looking at the screen, in total, per trial per register. Similarly to ManyBabies1 (The ManyBabies Consortium, 2020), a minimum looking time of 2 seconds was set as a criterion for inclusion of a trial in the data analysis.

## Planned statistical analyses

All data analysis will be conducted in the R Software (R Core Team, 2022) using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015), the *BayesFactor* package (Morey et al., 2015) and the *glmmTMB* to perform the weighted regression (Brooks et al., 2017), as well as the *PhonR* package for plotting of vowels in the vowel space area (McCloy, 2016a; McCloy, 2016b). The *ggbetweenstats* package (Patil, 2021) will also be used to visualize and explore the data. All p-values will be computed using the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017). Data preprocessing will also be performed in the R software (R Core Team, 2022).

The first hypothesis (H1a) and second hypothesis (H1b) will be assessed by performing a full-null model comparison to test for the potential effect of register (H1a) and paternity leave duration (H1b) and their possible interaction. The null model will contain each acoustic measure as a function of SES, while the full model also will contain register (IDS vs. ADS), paternity leave duration and their interaction:

Null model:

Acoustic measure ~ SES + (1+Register|Participant)

Full model:
Acoustic measure ~ Register*Pat_duration + SES + (1+Register|Participant)

For each acoustic measure, a separate comparison will be performed. A model test using the check_model() function from the *performance* package (Lüdecke et al., 2021) will be performed for model diagnostics and to visually check for various assumptions (normality of residuals, normality of random effects, linear relationship, homogeneity of variance, and multicollinearity); the acoustic measures would be transformed in cases of deviance from normality. In cases of an asymptotic (e.g. a sigmoid curve) relationship between paternity leave duration and acoustic measure, as would have been revealed by the function, a squared relationship will be added to the model.

To assess the third and the fourth hypotheses (H2a and H2b), the dependent variable will be the looking time (LT) in IDS and ADS trials. Here, H2a and H2b will be tested by performing a full-null model comparison to test for the potential effect of trial type (IDS vs. ADS), paternity leave duration and their possible interaction. The null model will contain looking time as a function of SES (maternal education), while the full model will also include the trial type, the paternity leave duration measure as well as their interaction:

Null model:
Looking time ~ SES + (1|subject)

Full model:
Looking time ~ Trial_Type*Pat_Duration + SES + (1|subject)

The model will be based on a data set with the data collapsed per child and register (IDS vs. ADS). In order to account for unequal sampling effort (i.e., variation in the number of successful trials per individual and register), we will weigh the contribution of the data points to the model by the number of completed trials per child and register. As a consequence, the more trials an infant has completed, the higher its contribution to the model. The model will be fitted with the function *glmmTMB* of the equally named package (Brooks et al., 2017), and the weights variable will be scaled such that the sum of the weights equals the total number of observations in the model. Prior to fitting the model we will z-transform Pat_duration and SES to a mean of zero and standard deviation of one to ease model convergence.

If the full-null model comparison is significant, we will inspect the individual predictors using the summary function on the model to determine what drives the effect(s).

**Exploratory analysis**

In order to assess the role of paternal attitudes and frequency of reading (how often the father read to his infant the past two weeks) on the acoustic measures of paternal IDS, we will conduct an exploratory analysis using a full-null model comparison approach for each acoustic measure. The null model will contain the aforementioned model for H1a, while the full model also will include paternal attitudes and reading as an activity.

Null model:

Acoustic measure ~ Register*Pat_duration + SES + (1+Register|Participant)

Full model:

Acoustic measure ~ Register*Pat_duration + SES + (1+Register|Participant) + Pat_attitudes + Pat_Reading

We will then perform a full-null comparison to test for the potential effect of paternal attitudes and reading as an activity. Here, we will compute the Variance Inflation Factor (VIF) to test for multicollinearity. If VIF>4, we will perform the full-null comparison for reading and paternal attitudes separately. If the effect of reading or paternal attitudes is significant, it will be used to generate a novel hypothesis for future research.

# References


Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi: 10.18637/jss.v067.i01

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.* Volume 36, Issue 4, December 2013, pp. 847-862. https://doi.org/10.1016/j.infbeh.2013.09.001

Benders, T., StGeorge, J., & Fletcher, R. (2021). Infant-directed speech by Dutch fathers: Increased pitch variability within and across utterances. *Language Learning and Development*, 17(3), 292-325. https://doi.org/10.1080/15475441.2021.1876698

Bergelson, E., Casillas, M., Soderstrom, M., Seidl, A., Warlaumont, A. S., & Amatuni, A. (2019). What do North American babies hear? A large-scale cross-corpus analysis. *Developmental Science*, 22(e12724), 1–12. https://doi.org/10.1111/desc.12724

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting Replicability in Developmental Research Through Meta-analyses: Insights From Language Acquisition Research. *Child Development*, *89*(6), 1996–2009. https://doi.org/10.1111/cdev.13079

Bingham, G. E., Kwon, K. A., & Jeon, H. J. (2013). Examining relations among mothers', fathers', and children's language use in a dyadic and triadic context. *Early Child Development and Care*, 183(3-4), 394-414.

Boersma, P. & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.12, retrieved 17 April 2022 from http://www.praat.org/

Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Maechler M, Bolker BM (2017). "glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling." *The R Journal*, 9(2), 378–400. https://journal.r-project.org/archive/2017/RJ-2017-066/index.html

Burnham, E. B., Wieland, E. A., Kondaurova, M. V., McAuley, J. D., Bergeson, T. R., and Dilley, L. C. (2015). Phonetic modification of vowel space in storybook speech to infants up to 2 years of age. *J. Speech Lang. Hear. Res.* 58, 241–253. doi: 10.1044/2015_JSLHR-S-13-0205

Byers-Heinlein, K., Bergmann, C., & Savalei, V. (2022). Six solutions for more reliable infant research. *Infant and Child Development*, e2296. http://doi.org/10.1002/icd.2296

Chládková, K., Escudero, P., & Boersma, P. (2011). Context-specific acoustic differences between Peruvian and Iberian Spanish vowels. *The Journal of the Acoustical Society of America*, 130(1), 416–428. https://doi.org/10.1121/1.3592242

Cristia, A. (2013). Input to Language: The Phonetics and Perception of Infant-Directed Speech. *Language & Linguistics Compass*, Volume 7, Issue 3, pp. 157-170. https://doi.org/10.1111/lnc3.12015

Cristia, A. & Seidl, A. (2014). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language*, *41(4)*, pp. 913 – 934. https://doi.org/10.1017/S0305000912000669

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child development*, 61(5), 1584-1595.

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

Englund, K. T., & Behne, D. M. (2005). Infant Directed Speech in Natural Interaction--Norwegian Vowel Quantity and Quality. *Journal of Psycholinguistic Research*, 34(3), 259–280. https://doi.org/10.1007/s10936-005-3640-7

Englund, K. & Behne, D. (2006) Changes in Infant Directed Speech in the First Six Months. *Infant and Child Development*, *15*, pp. 139-160.

Escudero, P., & Vasiliev, P. (2011). Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *The Journal of the Acoustical Society of America*, 130(5), https://doi.org/10.1121/1.3632043

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development*, 8(2), pp. 181–195. https://doi.org/10.1016/S0163-6383(85)80005-9

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

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

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.

Gergely, A., Faragó, T., Galambos, Á., & Topál, J. (2017). Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Scientific reports*, 7(1), 1-10.

Gleason, J. B. (1975). Fathers and other strangers: Men's speech to young children. *Developmental psycholinguistics: Theory and applications*, 1, 289-297.

Golinkoff, R. M., & Ames, G. J. (1979). A comparison of fathers' and mothers' speech with their young children. *Child development*, 28-32. https://doi.org/10.2307/1129037

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, 24(1), 14–20. https://doi.org/10.1037/0012-1649.24.1.14

Hayashi, A., Tamekawa, Y. & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language, and Hearing Research, 44*(6), pp. 1189-1200.

Hirst, D. (2022). «analyse_tier.praat» [Software]. Scripts and plugins for speech analysis with Praat. NAKALA. https://doi.org/10.34847/nkl.5df738jc

Hladik, E. G., & Edwards, H. T. (1984). A comparative analysis of mother-father speech in the naturalistic home environment. *Journal of psycholinguistic research*, 13(5), 321-332.

Jacobson, J. L., Boersma, D. C., Fields, R. B., & Olson, K. L. (1983).Paralinguistic features of adult speech to infants and small children.*Child Development*, 54, 436–442. https://doi.org/10.2307/1129704

Kalashnikova, M. & Burnham, D. (2018) Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *Journal of Child Language, Volume 45, Issue 5*, September 2018, pp. 1035 - 1053. https://doi.org/10.1017/S0305000917000629

Kartushina, N., & Mayor, J. (2022). Coping with dialects from birth: Role of variability on infants' early language development. Insights from Norwegian dialects. *Developmental Science*. https://doi.org/10.1111/desc.13264

Kuhl, P.K., Andruski, J. E., Christovich, I. A., Christovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyrovaulla, E. I., Sundberg, U., Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277,* 684-686. DOI: 10.1126/science.277.5326.684

Kuznetsova, A., Brockhoff, PB., Christensen, RHB (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1–26. doi: 10.18637/jss.v082.i13.

Le Chanu, M., & Marcos, H. (1994). Father–child and mother– child speech: A perspective on parental roles. European *Journal of Psychology of Education*, 9(1), 3–13.

Lennes, M. (2017). SpeCT - Speech Corpus Toolkit for Praat (v1.0.0). First release on GitHub. https://doi.org/10.5281/zenodo.375923

Leung, A., Tunkel, A. & Yurovsky, D. (2021) Parents Fine-Tune Their Speech to Children's Vocabulary Knowledge. *Psychological Science 2021*, Vol. 32(7), pp. 975–984. https://doi.org/10.1177/0956797621993104

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

Lüdecke et al., (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, 6(60), 3139, https://doi.org/10.21105/joss.03139

Ma, W., Golinkoff, R.M., Houston, D. M. & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development, 7(3)*, pp. 185-201. https://doi.org/10.1080/15475441.2011.579839

Marklund, E., & Gustavsson, L. (2020). The dynamics of vowel hypo-and hyperarticulation in Swedish infant-directed speech to 12-month-olds. *Frontiers in Communication*, 5, 523768.

McCloy, D. R. (2016a). phonR: tools for phoneticians and phonologists. R package version, 1.

McCloy, D. R. (2016b). Normalizing and plotting vowels with phonR 1.0.7. URL: https://drammock.github.io/phonR/

Morey, R. D., Rouder, Love, J., Marwick, B. (2015). BayesFactor package for R [software]. Computation of Bayes Factors for Common Designs. https://doi.org/10.5281/zenodo.31202

NAV (2022a). Hvor lenge kan du få foreldrepenger? Retrieved 2 June 2022 from: https://www.nav.no/foreldrepenger#hvor-lenge

NAV (2022b). Foreldrepenger 2021. Menn, dager, antall. Retrieved 23 May 2022 from https://www.nav.no/no/nav-og-samfunn/statistikk/familie-statistikk/tabeller/foreldrepenger-2021.menn-dager-antall

Newman, R. S., & Hussain, I. (2006). Changes in preference for infant-directed speech in low and moderate noise by 4.5-to 13-month-olds. *Infancy, 10*(1), 61-76.

OECD Family Database (2021). PF2.1. Parental leave systems. Organisation for Economic Co-Operation and Development. https://www.oecd.org/els/soc/PF2_1_Parental_leave_systems.pdf

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

Papoušek, M., Papoušek, H. & Haekel, M. Didactic adjustments in fathers' and mothers' speech to their 3-month-old infants. *J Psycholinguist Res* **16**, 491–516 (1987). https://doi.org/10.1007/BF01073274

Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach. *Journal of Open Source Software*, 6(61), 3167, doi:10.21105/joss.03167

Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior & Development*, 15(3), pp. 325–345. https://doi.org/10.1016/0163-6383(92)80003-D

Ferjan Ramírez, N. (2022). Fathers' infant-directed speech and its effects on child language development. *Language and Linguistics Compass*, 16(1), e12448.

Raneri, D., Von Holzen, K., Newman, R., & Bernstein Ratner, N. (2020). Change in maternal speech rate to preverbal infants over the first two years of life. *Journal of Child Language*, 47(6), 1263-1275. https://doi.org/10.1017/S030500091900093X

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rosslund, A., Mayor, J., Óturai, G., & Kartushina, N. (2022a). Parents' hyper-pitch and vowel category compactness in infant-directed speech are associated with 18-month-old toddlers' expressive vocabulary. *Language Development Research*. https://doi.org/10.31234/osf.io/wrku5

Rosslund, A., Mayor, J., Robberstad, S., Kartushina, N. (2022b). Mothers' and fathers' infant-directed speech have similar acoustic properties, and are associated with both direct and indirect measures of word comprehension in 8-month-old infants. Manuscript in preparation.

Rowe, M. L., Coker, D., & Pan, B. A. (2004). A comparison of fathers' and mothers' talk to toddlers in low-income families. *Social development*, 13(2), 278-291.

Schreiner, M. S., & Mani, N. (2017). Listen up! Developmental differences in the impact of IDS on speech segmentation. *Cognition*, 160, 98-102.

Schönbrodt, F. D. (2016). BFDA: Bayes factor design analysis package for R. https://github.com/nicebread/BFDA

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. Psychonomic bulletin & review, 25(1), 128-142. https://doi.org/10.3758/s13423-017-1230-y

Shapiro, N. T., Hippe, D., & Ferjan Ramírez, N. (2021). How chatty are daddies? An exploratory study of infants' language environments. Journal of Speech, Language, and Hearing Research, 64(8), 3242–3252. https://doi.org/10.1044/2021_JSLHR-20-00727

Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of Infant-Directed Speech on Early Word Recognition. *Infancy*, 14(6), pp. 654–666. https://doi.org/10.1080/15250000903263973

SR Research Ltd. (2022a). Data Viewer. Retrieved 11 May 2022 from https://www.sr-research.com/data-viewer/

SR Research Ltd. (2022b). Experiment Builder. Retrieved 11 May 2022 from https://www.sr-research.com/experiment-builder/

Steen, V. B., & Englund, N. (2022) Child-directed Speech in a Norwegian Kindergarten Setting, *Scandinavian Journal of Educational Research*, 66:3, 505-518, DOI: 10.1080/00313831.2021.1897873

The ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in Methods and Practices in Psychological Science 2020*, Vol. 3(1) pp. 24–52. https://doi.org/10.1177%2F2515245919900809

Thiessen, E. D., Hill E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, *7(1)*, pp. 53-71.

Weirich, M. & Simpson, A. (2019). Effects of Gender, Parental Role, and Time on Infant- and Adult-Directed Read and Spontaneous Speech. *Journal of Speech, Language and Hearing Research*, 62(11), 4001-4014. https://doi.org/10.1044/2019_JSLHR-S-19-0047

Weisleder, A., & Fernald, A. (2013) Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24,* pp. 2143-2152.

Xu Rattanasone, N., Burnham, D., and Reilly, R. G. (2013). Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age. *J. Phonet*. 41, 332-343. doi: 10.1016/j.wocn.2013.06.001