DOI: 10.1111/cdev.13950

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

Infant-directed speech does not always involve exaggerated vowel distinctions: Evidence from Danish

Christopher Cox[1,2,3] | **Christina Dideriksen**[1,2] | **Tamar Keren-Portnoy**[3] | **Andreas Roepstorff**[1,2] | **Morten H. Christiansen**[1,2,4] | **Riccardo Fusaroli**[1,2,5]

[1]School of Communication and Culture, Aarhus University, Aarhus, Denmark
[2]Interacting Minds Centre, Aarhus University, Aarhus, Denmark
[3]Department of Language and Linguistic Science, University of York, Vanbrugh College, York, UK
[4]Department of Psychology, Cornell University, New York, Ithaca, USA
[5]Linguistic Data Consortium, University of Pennsylvania, Philadelphia, USA

Correspondence
Christopher Cox, School of Communication and Culture, Aarhus University, Jens Chr. Skous Vej 2, Building 1485, 8200 Aarhus, Denmark.
Email: chris.mm.cox@gmail.com

Funding information
Danmarks Frie Forskningsfond, Grant/Award Number: DFF-7013-00074; Interacting Minds Centre

Abstract
This study compared the acoustic properties of 26 (100% female, 100% monolingual) Danish caregivers' spontaneous speech addressed to their 11- to 24-month-old infants (infant-directed speech, IDS) and an adult experimenter (adult-directed speech, ADS). The data were collected between 2016 and 2018 in Aarhus, Denmark. Prosodic properties of Danish IDS conformed to cross-linguistic patterns, with a higher pitch, greater pitch variability, and slower articulation rate than ADS. However, an acoustic analysis of vocalic properties revealed that Danish IDS had a reduced or similar vowel space, higher within-vowel variability, raised formants, and lower degree of vowel discriminability compared to ADS. None of the measures, except articulation rate, showed age-related differences. These results push for future research to conduct theory-driven comparisons across languages with distinct phonological systems.

INTRODUCTION

A mother sits on the floor with her 13-month-old infant. The mother points toward a cuddly toy and says in a high-pitched, animated voice: *Look! A penguin! Do you want to pet the penguin?*. The mother picks up a stuffed animal from the floor, holds it in front of the infant and repeats *Look! A penguin!* The infant then looks toward the mother and provides a rough attempt at repeating the word. The mother replies in an excited tone: *Yes, it's a penguin!*.

In the above scenario, we might expect the caregiver to address her child using a spontaneous form of language known as infant-directed speech (IDS). Across a diverse intersection of languages and cultures, the form of speech that adults direct to infants differs from that directed to other adults (i.e., adult-directed speech, ADS) in systematic ways (Cox, Bergmann, et al., 2022; Hilton et al., 2022). The acoustic characteristics of IDS have been studied extensively, and some clear commonalities have emerged across a wide variety of languages and cultures (e.g., Broesch & Bryant, 2015; Fernald et al., 1989; Gergely et al., 2017). For example, caregivers increase their vocal pitch and pitch variability, slow down their speech, and produce acoustically exaggerated vowels (e.g., Broesch & Bryant, 2015; Han et al., 2020; Kalashnikova et al., 2017; Kuhl et al., 1997). Many of the acoustic properties of IDS change as the infant becomes older in ways that allow caregivers to scaffold infants' early social and linguistic development (Cox,

Abbreviations: ADS, adult-directed speech; ALICE, Automatic LInguistic unit Count Estimator; elpd, expected log pointwise predictive density; IDS, infant-directed speech.

Bergmann, et al., 2022; Fusaroli et al., 2019; Fusaroli, Weed, et al., 2021; Warlaumont et al., 2014). This has led to the proposal that this speech style may serve various functions, such as allowing caregivers to express affect (cf. Benders, 2013), to regulate infants' emotional states (cf. Bryant & Barrett, 2007; Kitamura & Lam, 2009; Papoušek et al., 1991) as well as to facilitate language development (cf. Golinkoff et al., 2015; Kuhl, 2000). However, some questions about IDS require further investigation: (i) To what extent do common IDS properties generalize across a more diverse set of languages (cf. Blasi et al., 2022; Christiansen et al., 2022; Kidd & Garcia, 2022)? (ii) To what extent does the acoustic expression of IDS exhibit dynamic changes across the span of early development? (iii) Does IDS provide a clearer input than ADS, and how does this relate to the internal structure of vowel categories? In the remainder of this introduction, we detail each of these points and explain the aims of this comparative acoustic study of Danish IDS and ADS.

The extent to which the acoustic expression of IDS exhibits similar properties across languages with different phonological structures (e.g., Broesch & Bryant, 2015; Fernald et al., 1989; Han et al., 2020) and cultures with different parenting behaviors (e.g., Bergelson et al., 2019; Casillas et al., 2020) remains under-investigated. A recent meta-analysis of the acoustic properties of IDS found a high degree of between-study variability, part of which may arise due to differences in phonological systems across languages as well as methodological differences across studies of IDS (Cox, Bergmann, et al., 2022). Across these studies, there was a strong tendency for caregivers to produce IDS with a slower articulation rate as well as an elevated and more variable pitch across a wide variety of languages and cultures (Broesch & Bryant, 2015; Han et al., 2020; Narayan & McDermott, 2016). However, there were significant cross-linguistic differences in the degree to which caregivers produced acoustically exaggerated vowels (Cox, Bergmann, et al., 2022). Some languages, for example, showed a vowel space reduction in IDS (e.g., Benders, 2013; Englund & Behne, 2005; Rattanasone et al., 2013) and a higher degree of within-vowel variability (Cristia & Seidl, 2014; Martin et al., 2015; Miyazawa et al., 2017; Rosslund et al., 2022). The overrepresentation of North American English in all of these studies (Christiansen et al., 2022), which appears to provide a particularly exaggerated form of IDS (Fernald et al., 1989; Floccia et al., 2016), may inflate the importance of these findings. New data from a more diverse set of human cultures and languages (cf. Kidd & Garcia, 2022) can thus likely offer new insights into the possible forms of IDS in early infant development.

The current study focuses on Danish, a language characterized by a high degree of phonetic reduction, schwa assimilation, and a wide variety of vocalic sounds (cf. Basbøll, 2005; Højen & Nazzi, 2016; Trecca et al., 2021). This peculiar sound structure means that speakers of Danish often produce a speech stream with few or no clear spectral discontinuities to allow infants to engage in word segmentation (cf. Trecca et al., 2019, 2021). This in turn has led researchers to posit that Danish may be particularly difficult for infants to learn (Bleses et al., 2008a, 2008b, 2011), and even result in a vowel bias in word learning rather than a consonant bias (Højen & Nazzi, 2016). Looking-while-listening studies with Danish 2½-year-olds have shown that the vowel-rich nature of Danish has a negative impact on the processing of both known and novel words (Trecca et al., 2018, 2020). These findings provide potential explanations for the tendency for Danish infants to fall behind on some early linguistic milestones compared to children learning other languages (Bleses et al., 2008a, 2008b, 2011).

The peculiar sound structure of Danish raises the question as to how caregivers modify the acoustic properties of IDS. Two earlier investigations of IDS in Danish have painted a puzzling picture (Bohn, 2013; Dideriksen & Fusaroli, 2018). Both studies reported that IDS exhibited a slower articulation rate as well as an elevated f_o variability; however, the studies also reported either similar degree of vowel separability (Bohn, 2013) or even a slight reduction in vowel separability in IDS compared to ADS (Dideriksen & Fusaroli, 2018). The measure of between-vowel separability in these studies was derived by calculating the total area enclosed by steady-state formant frequencies of the three peripheral /i/-/a/-/u/ vowels, which has been posited to be a measure of speech clarity (Lam et al., 2012; Whitfield & Mehta, 2019), as discussed further below. The vowel inventory of Danish consists of 10 monophthongal vowel phonemes, seven of which are positioned in the front region of the vowel space (Grønnum, 1998). Because there are both short and long versions of each of these vowel phonemes—and only the long versions can be combined with the suprasegmental, creaky-voice feature of *stød*—the vowel inventory is estimated as comprising 30 phonologically distinct vowels, although estimates of this number differ depending on how the vowels are counted (Trecca et al., 2021). Given the high number of vowels that Danish-learning infants have to acquire, we might have expected caregivers (albeit unconsciously) to increase the separability between IDS vowel phonemes to provide a clearer speech input for their infants. However, these previous studies of Danish IDS provided a limited picture of its acoustic expression. First, Bohn (2013) presented results from a laboratory experiment with a high degree of experimental control. Parents were asked to talk about explicitly labeled toy animals to elicit specific words containing the /i/, /u/, and /a/ vowel tokens. Because the type of experimental task used to study IDS can influence the strength of the vowel space area effect (Cox, Bergmann, et al., 2022; cf. Marklund & Gustavsson, 2020), it is unclear whether this finding extends to spontaneous speech in Danish. Second, Dideriksen and Fusaroli (2018) used a coarse-grained approach to formant estimation by basing their data on automatic extractions of any type of voiced segment from the speech signal (cf. Degottex et al., 2014), which might unwittingly introduce biases in the estimates. To build upon the groundwork laid by these studies, we (i) relied on data collected in naturalistic settings, with caregivers and infants engaging in the types of interaction that take place in day-to-day activities, and (ii) conducted a detailed acoustic analysis of the internal distributions of each of the vowel categories in Danish IDS and ADS.

One of the benefits of IDS may derive from caregivers' dynamic adaptation to infants' developmental states (Fusaroli et al., 2019; Smith & Trainor, 2008) and responsiveness to infant vocalizations (Fusaroli, Weed, et al., 2023; Goldstein & Schwade, 2008; Ko et al., 2016; Nguyen et al., 2022; Warlaumont et al., 2014). Cross-linguistic studies of changes in the acoustic properties of IDS as a function of infant age, for example, indicated that caregivers increase their rate of articulation (Kondaurova et al., 2013; Lee et al., 2014; Narayan & McDermott, 2016; Raneri, 2015) and reduce the median f_o in IDS as their infants become older (Gergely et al., 2017; Han et al., 2020; Kondaurova et al., 2013; Niwano & Sugai, 2002; Vosoughi & Roy, 2012). Most studies on the properties of f_o variability and vowel space area, on the other hand, did not find evidence of any developmental shifts across a wide variety of age ranges (Benders, 2013; Burnham et al., 2014; Cristia & Seidl, 2014; Gergely et al., 2017; Hartman et al., 2017; Kalashnikova & Burnham, 2018; Lovcevic et al., 2020; Weirich & Simpson, 2019; Wieland et al., 2015). Two studies of vowel space area did show a shift over time, but contradicted each other in terms of the direction of the shift; whereas Japanese-speaking caregivers were reported to exhibit a gradual vowel space expansion as the infant became older (Dodane & Al-Tamimi, 2007), Cantonese Chinese-speaking caregivers were reported to reduce their vowel space as a function of infant age (Rattanasone et al., 2013). These acoustic modifications in IDS according to infants' preferences (Kitamura & Burnham, 1998; Singh et al., 2002) and processing limitations (Christiansen & Chater, 2016; Saffran & Kirkham, 2018) suggest that IDS involves interactive reciprocity, where active participation and reciprocity both play a crucial role (Goldstein & Schwade, 2008; Ko et al., 2016; Nguyen et al., 2022; Warlaumont et al., 2014). Investigating which aspects of the acoustic expression of IDS changes in the span of early infant development allows crucial insight into the different functions served by IDS (Bryant, 2022; Bryant & Barrett, 2007; Cox, Bergmann, et al., 2022). The question of how each of the acoustic features of Danish IDS undergo change with infant age will also be investigated in the current cross-sectional sample of participants.

A large body of research shows that the speech directed to infants has properties that facilitate speech processing (García-Sierra et al., 2016, 2021). For example, IDS stimuli can facilitate neural processing (Peter et al., 2016), induce faster word recognition (Song et al., 2010), and produce better word segmentation of an artificial language composed of nonsense syllables (Thiessen et al., 2005). The extent to which parents produce acoustically exaggerated vowels in IDS has been shown to have an effect on concurrent speech discrimination skills (García-Sierra et al., 2016, 2021; Liu et al., 2003), later expressive vocabulary (Hartman et al., 2017; Kalashnikova & Burnham, 2018), and the complexity of vocalizations at a later point in time (Marklund et al., 2021). Other descriptive studies showed positive correlations between pitch modulations in IDS and expressive vocabulary size (Porritt et al., 2014; Rosslund et al., 2022; Spinelli et al., 2017).

The facilitatory effects of IDS are generally attributed to its tendency to increase the clarity of the speech addressed to children (Hartman et al., 2017; Kuhl et al., 1997; Liu et al., 2003). Studies of adult speech perception have indicated a clear relation between vowel space expansion and speech intelligibility, the majority of which are conducted on American English (Bradlow et al., 1996, 2003; Ferguson & Kewley-Port, 2002, 2007; Lam et al., 2012; Liu et al., 2005; Whitfield & Goberman, 2017; Whitfield & Mehta, 2019). Computational models also learned vowel categories better if the location of category centroids was more distant from each another (De Boer & Kuhl, 2003; Eaves et al., 2016; McMurray et al., 2009; Vallabha et al., 2007). Other experiments showed that measures of speech clarity correlated with articulation rate and speech segment durations, suggesting that temporal aspects of speech may also play an important role (Ferguson & Kewley-Port, 2007; Lam & Tjaden, 2013; Searl & Evitts, 2013).

There is substantial evidence that vowel space areas tend to be larger in IDS than in ADS (e.g., Cristia & Seidl, 2014; Gergely et al., 2017; Kalashnikova & Burnham, 2018; Lam & Kitamura, 2012; Marklund & Gustavsson, 2020; Miyazawa et al., 2017; Weirich & Simpson, 2019); however, a number of languages, such as Dutch (Benders, 2013), Norwegian (Englund, 2018; Englund & Behne, 2005; Steen & Englund, 2021), and Cantonese (Rattanasone et al., 2013) show vowel space reduction in IDS. Although a lot of the literature on speech clarity is based on vowel space area, a number of authors have raised concerns about potential limitations to traditional vowel space measures (cf. Whitfield & Goberman, 2014, 2017). One of these limitations is that studies relying only on the acoustic measure of vowel space area disregard the crucial assumption that vowel space expansion in and of itself does not entail a greater degree of speech clarity. A number of studies have reported an increased degree of within-vowel variability in IDS (Cristia & Seidl, 2014; Martin et al., 2015; McMurray et al., 2013; Miyazawa et al., 2017; Rosslund et al., 2022). For example, Miyazawa et al. (2017) showed that Japanese mothers expanded the vowel space area of the IDS to their 18- to 20-month-old infants; however, due to an increase in within-vowel variability, this expansion did not lead to more distinct categories compared to those in ADS. It is thus only by computing the extent of within-vowel variability and between-vowel discriminability that claims about the clarity and overlap of phonetic categories can be made. The limitations of considering vowel space area alone thus makes it unclear whether the facilitative properties of IDS should be attributed to vowel space expansion, or whether this effect arises as an unintended side effect of another acoustic variable, such as vocal tract shortening through raising of the larynx (Kalashnikova et al., 2017) or smiling (Englund, 2018), or whether it simply occurs as a side effect of a slower articulation rate and therefore being able to reach articulatory targets (Whitfield & Goberman, 2017). To provide a fuller picture of the extent of within-vowel variability and between-vowel discriminability in Danish IDS and ADS, we manually annotated 9267 individual vowel tokens and analyzed the extent of compactness of individual vowel categories.

The present study was designed to provide a comprehensive analysis of the input from which Danish infants learn language by comparing the prosodic and vocalic properties of Danish caregivers' spontaneous IDS and ADS. To build on the insights from previous studies of Danish IDS (Bohn, 2013; Dideriksen & Fusaroli, 2018) and IDS across distinct languages (Cox, Bergmann, et al., 2022), we engaged in cumulative science practices and incorporated statistical results from these studies into our models to compare the findings. We turned our focus to the following three questions concerning the acoustic expression of Danish IDS and ADS:

(i) To what extent are five acoustic properties that have been reported extensively in the literature on IDS (Cox, Bergmann, et al., 2022) expressed in Danish ADS and IDS—fundamental frequency (f_o), f_o variability, articulation rate, vowel duration, and vowel space area?
(ii) Do any of the acoustic properties in Danish IDS exhibit age-related changes over the course of early infant development?
(iii) How do Danish parents negotiate the balance between within-vowel variability and between-vowel separability? Do vowel categories in IDS exhibit a higher degree of discriminability than in ADS?

METHODS

Participants and recording context

The speech recordings for this study consisted of spontaneous speech from 26 Danish-speaking mothers of infants between the ages of 11 and 24 months. The data overlapped somewhat with the dataset of a previous study of Danish IDS (Dideriksen & Fusaroli, 2018), which provided a coarse-grained analysis with a subset of the participants (N=10) in this dataset. Based on a precision analysis of a sample size of 26 participants (cf. Figures S5.1 and S5.2), repeated measures from this size of sample provided sufficient data to obtain credible intervals that allowed us to draw meaningful conclusions for the moderate effect sizes (Hedges' $g \simeq .5$) we expected from previous meta-analyses of the field (Cox, Bergmann, et al., 2022). The precision analysis indicated that—assuming a moderate effect size of 0.5—a within-subjects design with 26 participants allowed estimates with a standard deviation of 0.05. This was quite sufficient to draw robust conclusions about potential differences between the speech styles. The mothers spoke Danish as their first language, did not report any health problems, and their infants were not at risk for any linguistic or cognitive disabilities. All mothers provided informed consent. Each mother participated in two recording sessions: (i) to elicit spontaneous IDS, the mother was recorded for 60 min in a free play session with her infant, and (ii) to elicit spontaneous ADS, the mother participated in a casual conversation about summer holiday plans with an experimenter for 15 min. To increase the level of ecological validity, we took care to make the recording settings as natural as possible in order to capture the types of interaction that take place in day-to-day activities. To ensure that infants and caregivers felt comfortable, the recording sessions took place in participants' homes and used two video cameras (a Panasonic HDC-HS700 and GoPro Hero 5) and a high-quality wearable microphone. The experimenter interacted with the mother to elicit ADS, but remained absent during IDS recordings. Each mother was instructed to interact with her child as she would normally do in an everyday situation. The 26 infants (10 female) were aged between 11 and 24 months of age (mean = 16 months, SD = 4.4), as shown in Table S0.1.

Acoustic analysis

Diarization of the speech samples

In order to partition the IDS speech recordings into homogenous segments and identify utterance boundaries according to speaker identity, we used an open-source diarization tool known as ALICE (i.e., Automatic LInguistic unit Count Estimator; Räsänen et al., 2021). This software provides automatic diarization using a voice type classifier trained on over 200 h of day-long audio recordings of infants learning a number of typologically distinct languages (Räsänen et al., 2021). This voice type classifier uses the pyannote-audio library (Lavechin et al., 2020) and a neural network architecture called SincNet (Ravanelli & Bengio, 2018). Using this neural network architecture, the diarization tool extracts talker information from an input audio waveform and classifies utterances into talker categories. For the IDS recordings, we thus relied on this automatic diarization process and retained only segments spoken by the caregiver.

Because the ADS recordings involved two female interactants (i.e., the caregiver and experimenter), there would be a high risk of confusing the two voices in the automatic diarization process. We therefore transcribed all the audio files for the ADS speech samples and subsequently discarded utterances from the experimenter. An utterance was defined as a vocal production by one speaker that (i) was not interrupted by another speaker (e.g., backchannels were not considered) and (ii) did not contain a pause longer than a second. In most cases, the infant was in the room during the ADS recordings; however, interruptions from the infant were rare and all utterances from the infant as well as those directed toward the infant were discarded from the speech samples analyzed.

To ensure compatibility of the distinct diarization methods for the IDS and ADS recordings, we checked the accuracy of the ALICE output by comparing it to manually determined timecodes for 13 of the 26 recordings (cf. S10 in Supporting Information). Our analysis showed a substantial degree of agreement between the two methods for the timecodes (κ = .72 [.65; .84]) as well as substantial agreement between the extracted measures across the two types of diarization (cf. Figures S10.2 and S10.3).

Altogether, this diarization and data extraction process resulted in measures from 22,033 individual utterances: 3544 utterances in ADS and 18,489 utterances in IDS. The average length of utterance in ADS is more varied and on average includes longer utterances, as shown in Table 1. Although the number of utterances is smaller in ADS, the difference in mean length of utterance ensures that we have enough data on which to compare the speech styles, as shown by the comparable vocalic measures in Table 2.

TABLE 1 Information about length of utterances in the two speech styles.

Speech style	No. of utterances	Mean utterance length (s)	SD of utterance length (s)
ADS	3544	3.95	5.70
IDS	18,489	1.53	1.05

Abbreviations: ADS, adult-directed speech; IDS, infant-directed speech.

Extraction of prosodic measures

In order to extract the utterance-level acoustic data for this project, we wrote a custom Praat script based on the principles of De Jong and Wempe's (2009) script for syllable detection. We tested the script on a speech sample to obtain appropriate parameter values and ran the script with the same parameter values on all speech samples. The script detected potential syllables by extracting peaks of intensity above a threshold of 2dB above the median intensity of the utterance, with a time window of 64ms and a time step of 16ms. This captured the general tendency for a syllable nucleus to exhibit a higher degree of intensity than most surrounding sounds. The script then compared the intensity values of these potential syllables to the preceding dips in intensity and disregarded intensity peaks that do not show a preceding dip of at least 2dB with respect to the current peak. By dividing the total number of syllables with the duration of the individual utterances, we calculated the articulation rate for each utterance. At the same time, the script extracted the pitch contour of the utterance using a window size of 100ms and time step of 20ms. The script then summarized the data in terms of median f_o (centralized measure of f_o in Hz), interquartile range of f_o (in Hz), as well as articulation rate (syllables per second). We thus had 22,033 individual utterance-level data points for each of these prosodic measures. We should note that these acoustic properties exhibit interdependence, as shown by the correlation network plots for each of the speech styles in Figure S4.1. Moreover, the process of automatically extracting these acoustic measurements from the audio may have generated some errors. To combat the influence of these potential measurement errors, we chose to use robust regression over outlier detection and deletion for the following two reasons: (i) it is difficult if not impossible to establish an objective definition of what counts as a measurement error versus what is an inherent property of the distribution, and (ii) outlier deletion wastes data and can lead to underestimation of the variance (Yellowlees et al., 2016). We ran an additional check to determine whether this choice influenced our estimates by comparing the results of models with and without outlier deletion (i.e.,

TABLE 2 Overview of the number of vowels in each quantile of vowel duration as well as across phonological length for each speech style. The quantiles of duration are based on the entire dataset of vowels.

	No. of vowels <20%	No. of vowels 20%–40%	No. of vowels 40%–60%	No. of vowels 60%–80%	No. of vowels 80%–100%	No. of long vowels	No. of short vowels	Total no. of vowels
ADS	877	934	908	889	917	390	4133	4523
IDS	983	913	947	965	937	505	4239	4744

Abbreviations: ADS, adult-directed speech; IDS, infant-directed speech.

values above and below two standard deviations from the mean). As shown in Table S1.5, the control analysis of each of the utterance-level acoustic variables indicated similar model estimates and credible intervals for the data with and without outlier deletion.

Extraction of vocalic measures

To analyze the duration and formants of individual vowel categories in Danish across the two speech styles, we manually segmented the onsets and offsets of 9267 vowels using Praat (Boersma & Weenink, 2022). The vowel tokens were not annotated if the acoustic signal was disrupted by ambient noise or overlapping speech, if the vowel was too short to identify a stable midpoint, or if clear formants were not present due to whispered speech or creaky voice (including *stød*). For each speaker, we annotated any vowel that conformed to the above criteria until we had sampled a minimum of 150 vowel tokens in each speech style. The total number of vowels ended up being slightly higher than 7800 due to our exploration of how to extract the vowels in the initial phases of the analysis process; that is, for subjects AF, AN, and CL we have 488, 792, and 486 vowel tokens, respectively, while the number of vowel tokens for the rest of the subjects are between 307 and 348. This sampling approach to the spontaneous speech data allows us to gain insight into the vowel types produced by each speaker while limiting degrees of freedom in the analysis process. To perform control analyses of the vowel data, moreover, we annotated each vowel according to the following four binary properties: phonological length (long vs. short), stress (stressed vs. unstressed), focus (whether the vowel appears in a focused constituent in an utterance), and word type (content word vs. function word). We wrote a custom Praat script to extract the first three formants at the temporal midpoint of each vocalic interval, vowel duration as well as the above four binary properties. We should note that the short versions of four Danish vowels, notably /ɛ/, /œ/, /ɔ/, and /u/, have been shown to be slightly more centralized than the long versions in the Aarhus dialect (Steinlen, 2005). We assumed that the ratio of vowels with short and long durations remained similar in the two speech styles, thereby not influencing the between-style comparison of vowel quality. To check this assumption, we computed the number of vowels in each of the speech styles across different quantiles of vowel duration. As shown in Table 2, we see similar distributions of the number of vowels across quantiles, phonological length, and in the total number of vowels analyzed.

A random sample of three spectrograms for each of the seven border vowels in Danish IDS from one of the speakers is shown in Figure 1. The script considered the maximum value for the formant search to be 5500 Hz and applied pre-emphasis to frequencies above 50 Hz. The raw vowel formant values for each vowel category are provided in Table 3.

These vowel formant data were then imported and further analyzed in R Studio (RStudio Team, 2020). To minimize the acoustic variation that arises due to physiological and anatomical differences across speakers, the vowel formant data were normalized using the vowel-extrinsic methods of Lobanov's (1971) z-score transformation. This normalization procedure has been shown to reduce inter-speaker anatomical/physiological variation while preserving phonemic variation, without requiring logarithmic scale transformation (Adank et al., 2004; Nearey, 1978). The procedure involved subtracting the mean speaker-specific formant value from all vowel tokens for a given speaker, and then dividing by the speaker-specific standard deviation for that formant. This procedure expresses formant values according to the hypothetical center of each speaker's vowel space and was developed for automatic speech recognition purposes (Lobanov, 1971). The rest of the results in this paper are reported using the normalized vowel formant data. The raw formant data and code have been made available on OSF: https://osf.io/ywf9m/?view_only=d99fc6dbc61546febff619b8674a7943.

Quantification of vowel space area, vowel variability, and distinctiveness

We used the phonR package (McCloy, 2016) to calculate two measures of vowel space area for each subject and speech style: (i) the vowel space enclosed by plotting the median z-score-transformed first and second formant values of /i/, /a/, and /u/ in a Cartesian coordinate system, as in traditional studies of vowel space area (Kuhl et al., 1997; Liu et al., 2003), and (ii) the vowel space encompassed by the median z-score-transformed formant values of all of the border vowels Danish (i.e., /i/, /e/, /ɛ/, /ɑ/, /ʌ/, /ɔ/, /o/, and /u/). This latter measure was computed in order to be able to assess the difference between the two speech styles in terms of the total area encompassed by all of the vowels. To avoid confusion between these two measures in the paper, we henceforth use vowel space area to refer to the traditional /i/−/a/−/u/ measure (cf. Kuhl et al., 1997; Liu et al., 2003) and refer the reader to Section S3 whenever we discuss the vowel space area results based on all eight border vowels. The calculation of each of the vowel space area measures is conducted for each subject in each speech style. The measure thus consists of 52 data points (i.e., 26 for each speech style). To facilitate interpretability of this measure, moreover, we z-transformed the vowel space areas by subtracting the mean vowel space area and dividing by the standard deviation. This rescales the vowel space area to be on the same scale as the hedges' g standardized effect size; a positive score on this scale indicates vowel space expansion while a negative score indicates vowel space

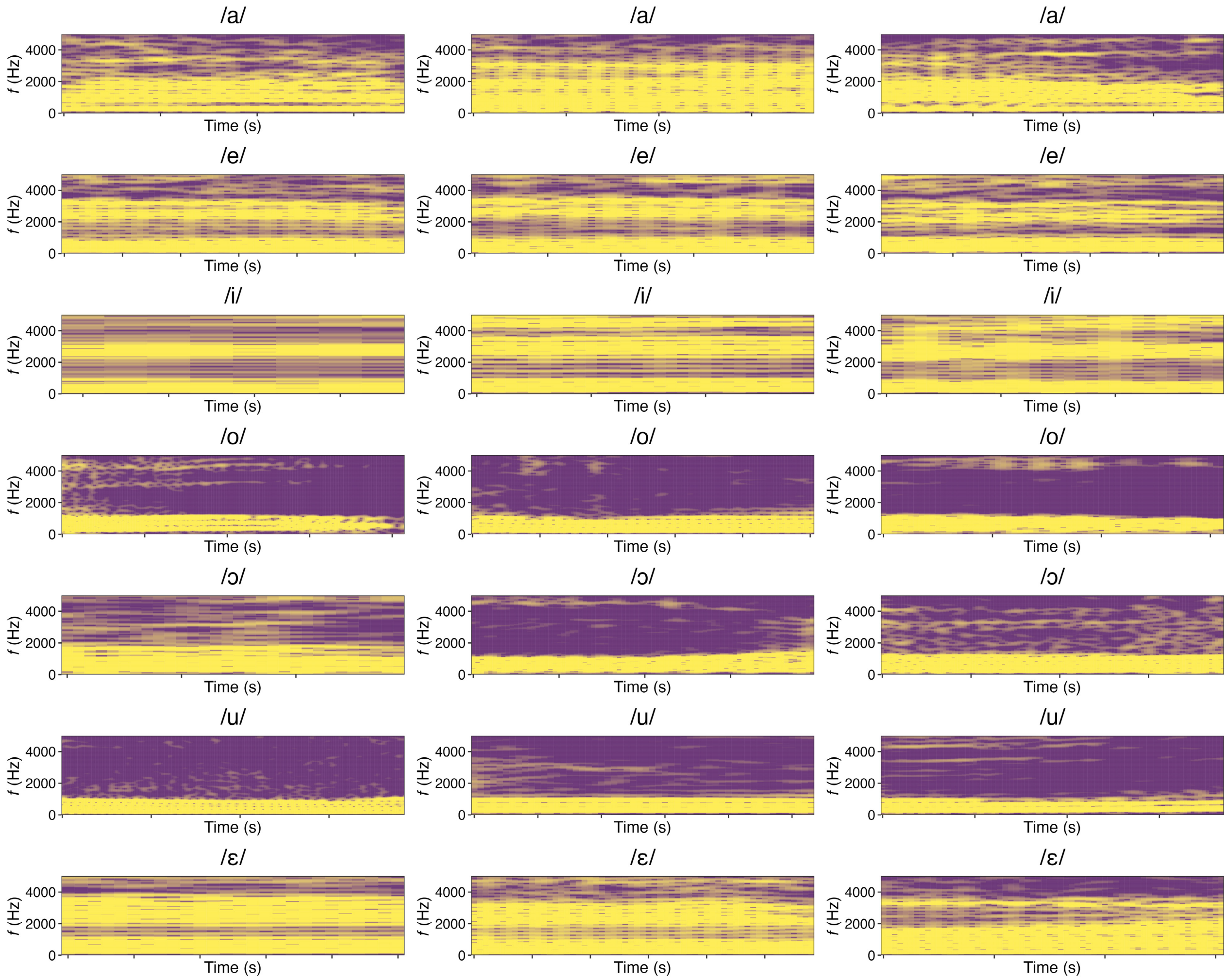

FIGURE 1 Three random samples of vowel tokens for seven of the peripheral vowels analyzed. The more yellow the spectrogram, the higher the amplitude of the frequency present in the audio signal. The limits of the *x*-axis show the total time for each of the vowel tokens displayed.

reduction. We also conducted control analyses for vowel space areas based on the border vowels according to the binary properties noted for each of the vowels: stress, focus, length, and word type, as shown in Figure S6.2.

Furthermore, we use the extracted formant data to estimate the extent of variability in each of the two formant dimensions for each vowel category across the two speech styles. To compare the extent of variability within each vowel category across the speech styles, we quantified the evidence in favor of a greater degree of variability in formant values within each speaker, as explained further in our choice of statistical models in the next section. Similarly, we adapted Rosslund et al.'s (2022) approach to calculating the distinctiveness of vowel categories in the two speech styles. While vowel variability indicates the compactness of vowel tokens within each category, vowel discriminability quantifies the degree of overlap between categories, which has also been argued to be a measure of speech intelligibility (cf. Kim et al., 2011; Whitfield & Goberman, 2014, 2017). We computed a measure of vowel discriminability for each participant within each speech style as the squared distances of category centroids from the overall vowel space centroid, divided by the squared distances of individual vowel tokens from the overall vowel space centroid. This continuous measure thus indexed the proportion of variance in F1 and F2 explained by the vowel category identity: a value of 1 would indicate that vowel membership explained all variance (i.e., a low degree of overlap between categories) and 0 would indicate that vowel membership explained no variance (i.e., a high degree of overlap between categories). Due to its reliance on the squared distances of vowel tokens, this measure of distinctiveness can thus be construed as the proportion of variance explained by category membership; the more distinctive the vowel categories, the higher the explained variance due to category membership (cf. Rosslund et al., 2022).

TABLE 3 Number of tokens (*n*), median duration (Dur) in seconds, median first formant (F1), and second formant (F2) in Hz as well as their standard deviations (SD) for each target vowel across IDS and ADS speech styles. It should be noted that the numbers in this table refer to the non-normalized formant values for each of the vowel categories. See Figures S7.1–S7.4 and Tables S8.1–S8.4 to see the raw formant values for each speech style across different contexts of word type, phonological length, focus, and stress.

ADS					IDS				
Vowel	*n*	Dur (SD)	F1 (SD)	F2 (SD)	Vowel	*n*	Dur (SD)	F1 (SD)	F2 (SD)
æ	158	0.097 (0.06)	651.9 (114.7)	1879.3 (240.8)	æ	148	0.087 (0.04)	696.3 (111.4)	1888.6 (325.9)
ɑ	1040	0.083 (0.04)	792.4 (139.4)	1598 (204.4)	ɑ	968	0.08 (0.05)	792.7 (157.5)	1711.1 (264.2)
e	674	0.072 (0.04)	483.5 (85.8)	2052.9 (245.7)	e	724	0.071 (0.04)	534.7 (114.2)	2067.6 (384.3)
i	391	0.066 (0.03)	424.4 (79.7)	2201.4 (350)	i	305	0.076 (0.05)	455.7 (100.3)	2217.5 (473)
o	92	0.079 (0.04)	502.5 (105)	1069.1 (207.8)	o	179	0.088 (0.04)	591.2 (108.8)	1145.5 (240.8)
ø	87	0.071 (0.03)	472.8 (77.6)	1708 (160.9)	ø	105	0.076 (0.06)	580.1 (129)	1875.2 (223.5)
œ	6	0.075 (0.02)	529.6 (31.7)	1619.2 (175.4)	œ	31	0.077 (0.04)	617.5 (115)	1769.1 (338.3)
ɶ	103	0.093 (0.03)	688.9 (106.4)	1593.9 (136.9)	ɶ	54	0.088 (0.05)	688.5 (97.6)	1634.9 (157)
ɔ	181	0.08 (0.04)	533.2 (87.8)	1220.6 (205.8)	ɔ	336	0.089 (0.06)	641.2 (115.1)	1318.3 (202.1)
u	101	0.063 (0.05)	448.4 (91.4)	1152 (232.1)	u	157	0.068 (0.07)	487.8 (115.6)	1236.4 (311)
ʌ	863	0.08 (0.06)	635.7 (99.4)	1427.5 (175.7)	ʌ	691	0.079 (0.05)	726.2 (128.4)	1491.7 (210)
y	56	0.066 (0.03)	423.3 (83.6)	1893.8 (290.8)	y	65	0.071 (0.03)	467.6 (94.1)	1998.1 (221.5)
ɛ	576	0.072 (0.04)	581.3 (111.7)	1877.3 (208.5)	ɛ	781	0.073 (0.04)	655.1 (135.5)	1896 (297)

Abbreviations: ADS, adult-directed speech; IDS, infant-directed speech.

Statistical modeling

To assess the extent to which the acoustic properties of Danish ADS and IDS differ, we ran Bayesian multi-level robust regression models of the data. For all of the measures appearing only on a positive range of values (i.e., all measures except vowel space area), we used a logarithmic link function to model the potential long tails of high values in the data (cf. Gabry et al., 2019; McElreath, 2018), as explained further in Section S1. For each acoustic measure, we built three models: (i) an intercepts-only model with varying effects by participant, (ii) a model with speech style (i.e., IDS vs. ADS) as a predictor with varying subjects nested within speech style, and (iii) a model with an interaction term between speech style and age as a predictor, as well as varying intercepts and slopes for subjects nested within speech style. In the models with speech style as a predictor (i.e., model (ii) and (iii)), we allowed the model to estimate a separate sigma across the two speech styles; that is, a different expected error when predicting data in the two speech styles (i.e., heteroskedasticity). We made this choice to explore the potential effects of a slight imbalance in the number of utterance-level measures for each of the speech styles (cf. Table 1) and because we might expect a greater amount of heterogeneity in caregivers' interactions with children (Englund, 2018).

To model the location (centroid) and scale (variability) of each of the vowel categories under investigation, moreover, we built hierarchical mixed-effects location-scale models for the *z*-score-transformed F1 and F2 measures. This type of location-scale model represents an extension to multi-level regression models in that they allow estimation of covariances among the random effects, both within and across the location and scale of the predictor (Hedeker et al., 2008; Williams et al., 2019). This form of model offers new insights into the structure of individual vowel categories, estimation of intra-vowel variability, extent of formant raising, and the influence of speech style upon these parameters (Rast & Ferrer, 2018). We should note that in our location-scale models of vowel categories, we restrict our analysis to vowel categories for which have above 50 tokens (cf. Table 3) to avoid biasing the estimates.

All computations were performed in R 4.2.0 (R Core Team, 2020) using *brms* 2.16 (Bürkner, 2017) and *cmdstanr* 2.28.2 (Carpenter et al., 2017) in RStudio 1.4 (RStudio Team, 2020). We report the model formulae, model specifications, and our choice of priors in Section S1.1. We chose weakly informative priors in order to discount extreme values as unlikely and to ensure minimal influence on the posterior estimates of the model (Gelman et al., 2017; Lemoine, 2019). We provide explicit description and visualization of our choice of priors, as well as prior sensitivity checks in Sections S1.2–S1.4.

Throughout the rest of this paper, we provide estimates and report 95% credible intervals, evidence ratios, and credibility scores. The credible interval refers to the range of values within which there is a 95% probability that the true parameter value lies given the assumptions of the model. We report these intervals in square brackets. The evidence ratio denotes the ratio of likelihood in favor of a hypothesis. An evidence ratio of 10, then, implies that the hypothesis is 10 times more likely than the alternative. An evidence ratio of "Inf" (infinite) occurs when all the posterior samples conform to the direction of the hypothesis and not to alternative directions. The credibility score refers to the proportion of posterior samples in the direction of the hypothesis under investigation. We perform leave-one-out information criterion-based model comparison (Vehtari et al., 2017) between the three models and report the differences in expected log pointwise predictive density difference (elpd). This measure refers to the difference between the models in terms of their expected out-of-sample predictive accuracy; that is, it quantifies the extent to which model predictions generalize to an independent dataset. The lower the out-of-sample predictive accuracy, the lower the elpd (Vehtari et al., 2017; Yao et al., 2018).

RESULTS

The results are structured as follows. We start by reporting results that concern the first aim of the study, namely to quantify five acoustic properties of Danish IDS: (i) median f_o, (ii) f_o variability, (iii) articulation rate, (iv) vowel duration, and (v) vowel space area. We compare these results to those of a previous study on Danish IDS (Bohn, 2013) as well as a recent large-scale meta-analysis of these same acoustic features (Cox, Bergmann, et al., 2022) by incorporating prior statistical information into our models. We then present results pertaining to our second aim, namely to investigate the extent to which the acoustic measures exhibit age-related change. We restrict ourselves to depicting the acoustic measure as a function of age only if age exhibits robust effects (i.e., only for articulation rate). We refer the reader to the age plots in Section S2. We then present results that concern the third aim, namely to quantify the properties of within-vowel variability and between-vowel discriminability across the two speech styles.

Median f_o

The data for median f_o indicate that caregivers' utterances in IDS (235.8 Hz with 95% CI [228.5; 243.2]) exhibit a higher median f_o than in ADS (202.9 Hz with 95% CI [196.0; 209.6]), as shown in Figure 2. We obtain strong evidence in favor of the hypothesis that IDS exhibits a higher f_o (33.0 Hz [27.9; 38.3], evidence ratio: Inf, credibility score: 1, with 26/26 participants displaying a positive effect). The model indicates a greater amount of variability in the distribution of overall median f_o values in IDS (σ=45.2 Hz [43.1; 47.5]) than in ADS (σ=23.7 Hz [21.0; 26.7]). We accordingly also obtain strong evidence in favor of the hypothesis that IDS exhibits a greater amount of residual variability within speech style (21.5 Hz [19.0; 24.01], evidence ratio: Inf, credibility score: 1, with 26/26 participants displaying a positive effect). The model indicates a robust correlation between IDS and ADS values (r=.66 [.38; .84]). This correlation implies that there is a clear tendency for the height of subjects' f_o in IDS to depend on their corresponding f_o value in ADS. The model with age as a predictor exhibits less out-of-sample predictive accuracy (elpd: −795.3, se: 44.8) than the model without age, and we accordingly see no robust effect of age (estimate=.002 [−.003; .008], evidence ratio=3.36, credibility score=0.77), as shown by Figure S2.

f_o variability

The data for the interquartile range of f_o indicate that caregivers produce a higher degree of f_o variability in IDS (74.9 Hz [71.1; 79.0]) than in ADS (48.9 Hz [44.3; 53.8]), as shown in Figure 3. Similar to median f_o, we obtain strong evidence in favor of the hypothesis that f_o variability is higher in IDS than in ADS (26 Hz [21.8; 30.2], evidence ratio: Inf, credibility score: 1, with 26/26 participants displaying a positive effect). The model indicates a greater amount of variability in the distribution of overall f_o variability values for IDS (σ=44.3 Hz [42.7; 46.0]) than for ADS (27.1 Hz [23.9; 30.7]). We also obtain strong evidence in favor of the hypothesis that IDS exhibits a greater amount of residual variability in overall f_o variability values within speech style (17.1 Hz [14.4; 19.7], evidence ratio: Inf, credibility score: 1, with 26/26 participants displaying a positive effect). The model finds weak evidence for a correlation between the speech style estimates for each subject (.35 [−.02; .65]). The model with age as a predictor exhibits less out-of-sample predictive accuracy (elpd: −530.5, se: 38.2) compared to the model with only speech style as a predictor, and we accordingly find no robust effect of age (estimate=.000 [−.011; .010],

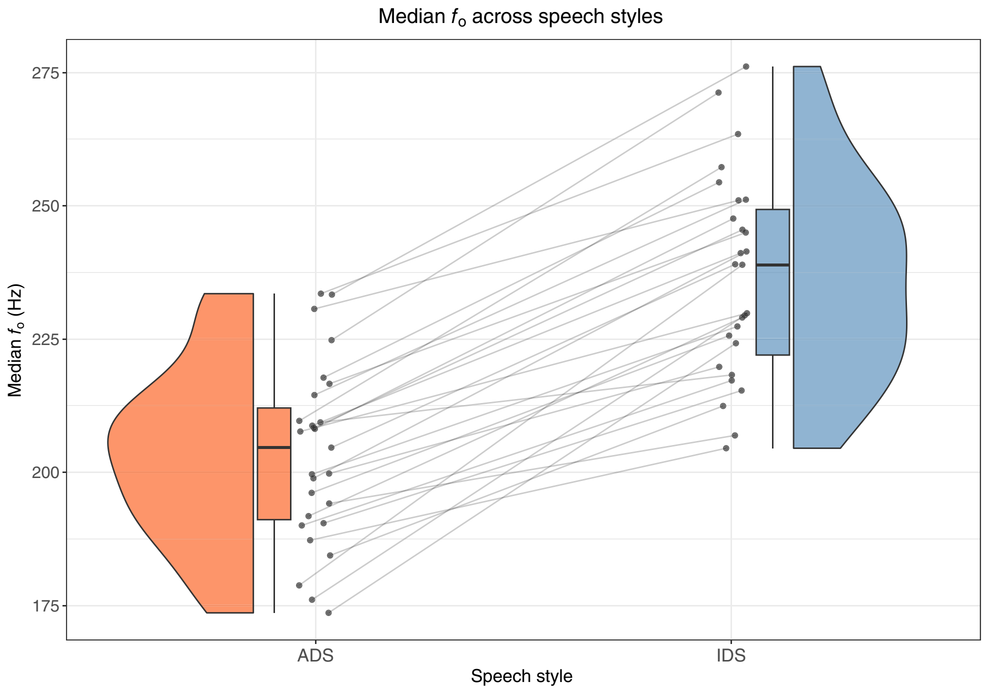

FIGURE 2 Plot of model estimates for individual subjects' median f_o across the two speech styles. Each point in each speech style indicates one subject. The points for each subject are connected across the two speech styles with a line. Because model estimates are based on a pooling of repeated measures for each participant, we plot these estimates to provide a more robust picture of differences across speech styles. ADS, adult-directed speech; IDS, infant-directed speech.

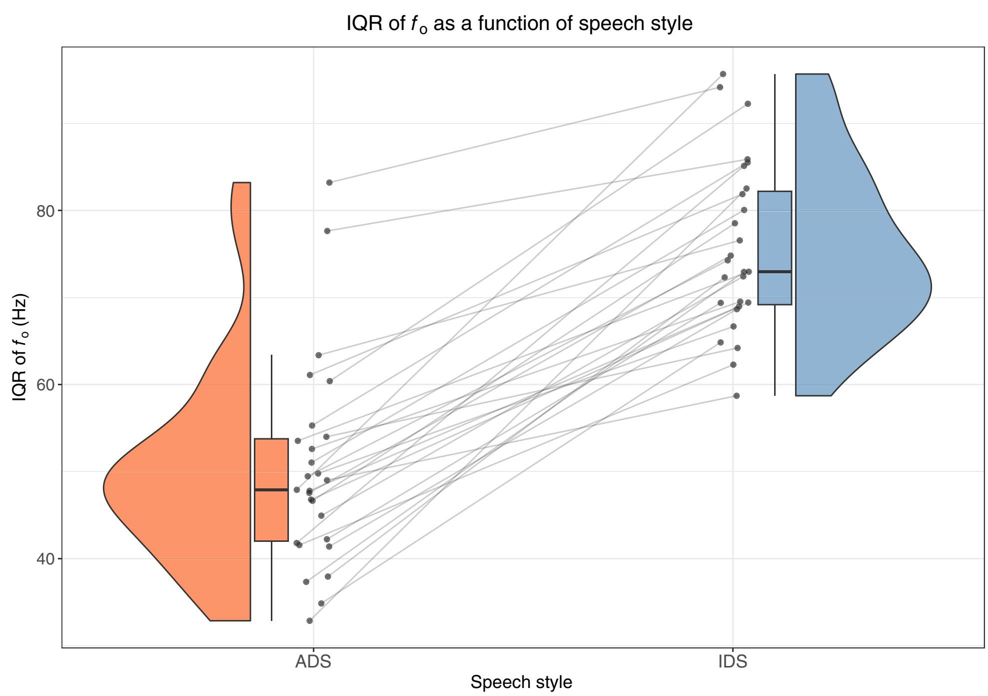

FIGURE 3 Plot of model estimates for subjects' pitch variability across the two speech styles. Each point in each speech style indicates one subject. The points for each subject across the two speech styles are connected with a line. Because model estimates are based on a pooling of repeated measures for each participant, we plot these estimates to provide a more robust picture of differences across speech styles. ADS, adult-directed speech; IDS, infant-directed speech.

evidence ratio = 0.85, credibility score = 0.46), as shown by Figure S2.

Articulation rate

The data for articulation rate indicate that caregivers produce fewer syllables per second in IDS (2.96 syll/s with 95% CI [2.89; 3.02]) compared to ADS (3.72 syll/s with 95% CI [3.61; 3.84]), as shown in Figure 4. We obtain strong evidence in favor of the hypothesis of a slower articulation rate in IDS (−0.77 syll/s [−0.87; −0.67], evidence ratio: Inf, credibility score: 1, with 26/26 participants displaying a negative effect). The model indicates a higher degree of variability in IDS (σ = 1.1 syll/s [0.958; 1.06]) than in ADS (σ = 1.07 syll/s [1.06; 1.09]). We accordingly obtain evidence that articulation rate in IDS exhibits a greater amount of residual variability within speech style (0.06 syll/s [0.02; 0.11], evidence ratio = 139.6, credibility score = 0.99, with 26/26 participants displaying a positive effect). The model indicates no clear correlation between the speech style estimates for each subject (r = .11 [−.29; .49]). The model with age as a predictor exhibits similar out-of-sample predictive accuracy (elpd: −23.4, se: 6.9) compared to the model with only speech style as a predictor, and we observe a robust small effect of age (estimate = .006 [.002; .009], evidence ratio = 172.1, credibility score = 0.994), as shown in Figure 5 as well as Figure S2.

Vowel duration

The data for vowel duration indicate that caregivers overall produce similar vowel durations in ADS (0.078 s [0.075; 0.081]) and in IDS (0.077 s [0.074;0.081]), as shown in Figure 6. We obtain no evidence in favor of the hypothesis that caregivers' vowels in IDS exhibit longer duration (−0.000 s [−0.004; 0.003], evidence ratio = 0.61, credibility score = 0.38, with only 6/26 participants displaying a reliable positive effect). The model also indicates a similar degree of variability in the distribution of values in IDS (σ = 0.024 [0.022; 0.025]) and ADS (σ = 0.024 [0.022; 0.026]). We accordingly obtain no evidence in favor of the hypothesis that IDS exhibits a greater amount of residual variability within speech style (0.000 s [−0.001; 0.002], evidence ratio: 1.33, credibility score: 0.57, with only 5/26 participants showing a reliable positive effect). The model with age as a predictor exhibits less out-of-sample predictive accuracy (elpd: −77.7, se: 14.5) compared to the model without age as a predictor, and we likewise see no robust effect of age (estimate = .002 [−.007; .010], evidence ratio = 1.80, credibility score = 0.64), as shown by Figure S2. We conducted a control analysis of vowel duration to check whether the speech styles exhibited different values across phonologically short and long vowels (cf. Figure S6.1). The control analysis indicated that short vowels are slightly longer in ADS than in IDS, however, not robustly so (estimate = .02 [−.02; .06], evidence ratio = 3.1, credibility

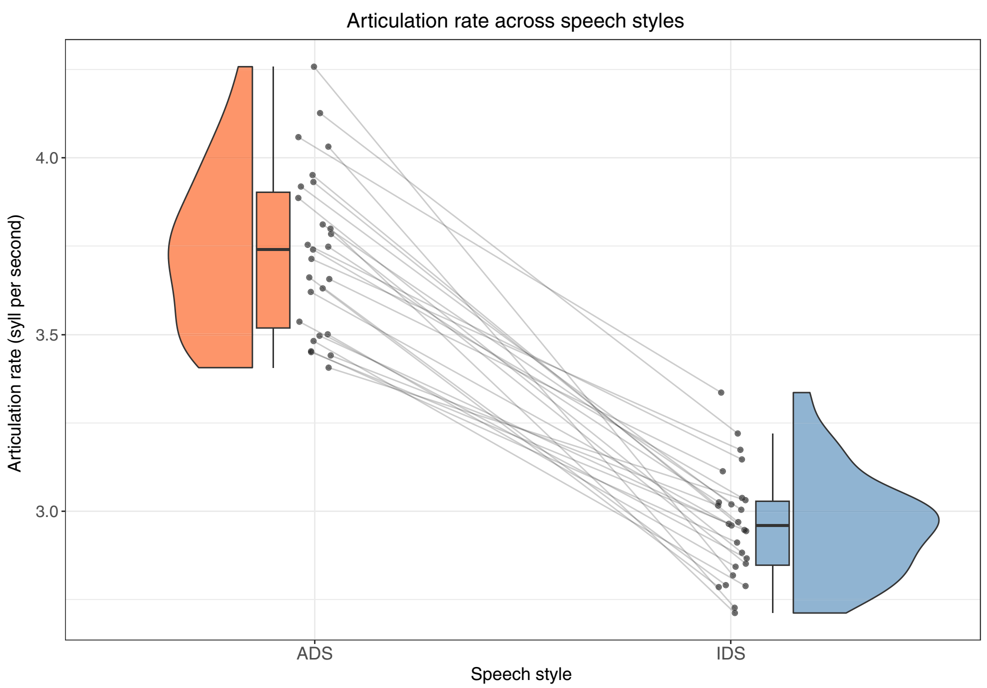

FIGURE 4 Plot of model estimates for subjects' articulation rate across the two speech styles. Each point in each speech style indicates one subject. The points for each subject are connected across the two speech styles with a line. Because model estimates are based on a pooling of repeated measures for each participant, we plot these estimates to provide a more robust picture of differences across speech styles. ADS, adult-directed speech; IDS, infant-directed speech.

score = 0.77), whereas long vowels in ADS were shown to be longer than those in IDS (estimate = .07 [.01; .13], evidence ratio = 30.5, credibility score = 0.97). We conducted a second control analysis by normalizing vowel duration according to the inverse of the median articulation rate for each speaker (cf. Figures S9.1 and S9.2). This control analysis indicated that rate-normalized vowel length in IDS (0.23 [0.22; 0.24]) was robustly smaller than in ADS (0.30 [0.29; 0.31]).

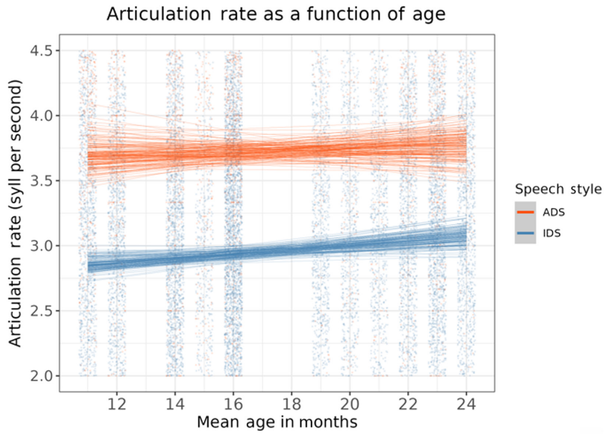

FIGURE 5 Plot showing how caregivers' articulation rate changes as a function of infant age in the two speech styles. The plot shows 150 posterior predictions from the model including age as a predictor. The plot indicates relative stability for ADS and a relative increase in articulation rate in IDS. ADS, adult-directed speech; IDS, infant-directed speech.

Vowel space area

The data for vowel space area indicated that caregivers' vowel space area in ADS (Cohen's $d = -.01$ [−.37; .34]) was greater than that in IDS (Cohen's $d = -.44$ [−.81; .05]). As shown in Figure 7, we obtained strong evidence in favor of the hypothesis that the vowel space area in IDS was smaller than in ADS ($d = -.43$ [−.85; −.001], evidence ratio = 19.18, credibility score = 0.95, however, the posterior estimates for all 26 participants included the null within their credible interval). The model indicated similar degree of substantial variability in ADS ($\sigma = 0.88$ [0.52; 1.38]) and IDS ($\sigma = 0.94$ [0.55; 1.49]). We obtained no evidence that IDS exhibited a greater degree of residual variability in vowel space area estimates within speech style ($d = .09$ [−.43; .62], evidence ratio = 1.58, credibility score = 0.61, with 0/26 participants displaying an effect without the null in their credible interval). The model with age as a predictor exhibited similar out-of-sample predictive accuracy (elpd: −1.3, se: 2.2) compared to the model without age as a predictor; however, as shown in Figure S2, the measure exhibited a high degree of uncertainty (estimate = −.04 [−.07; −.01], evidence ratio = 31.64, credibility score = 0.969). These results should be interpreted with caution, especially as the model of vowel space

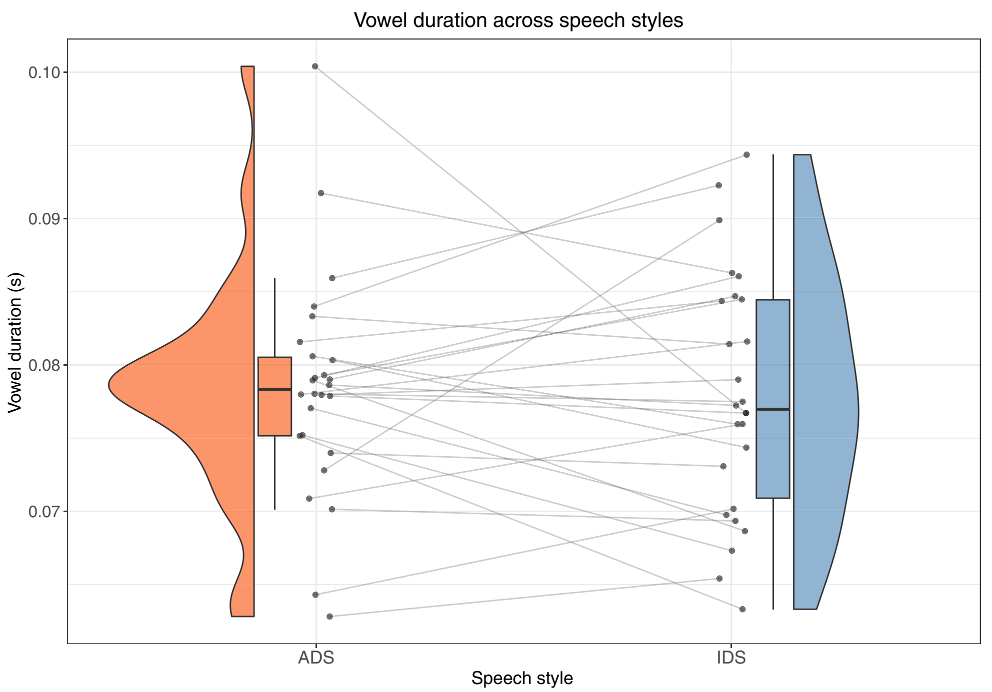

FIGURE 6 Plot of model estimates for subjects' vowel duration across the two speech styles. Each point in each speech style indicates one subject. The points for each subject are connected across the two speech styles with a line. Because model estimates are based on a pooling of repeated measures for each participant, we plot these estimates to provide a more robust picture of differences across speech styles. ADS, adult-directed speech; IDS, infant-directed speech.

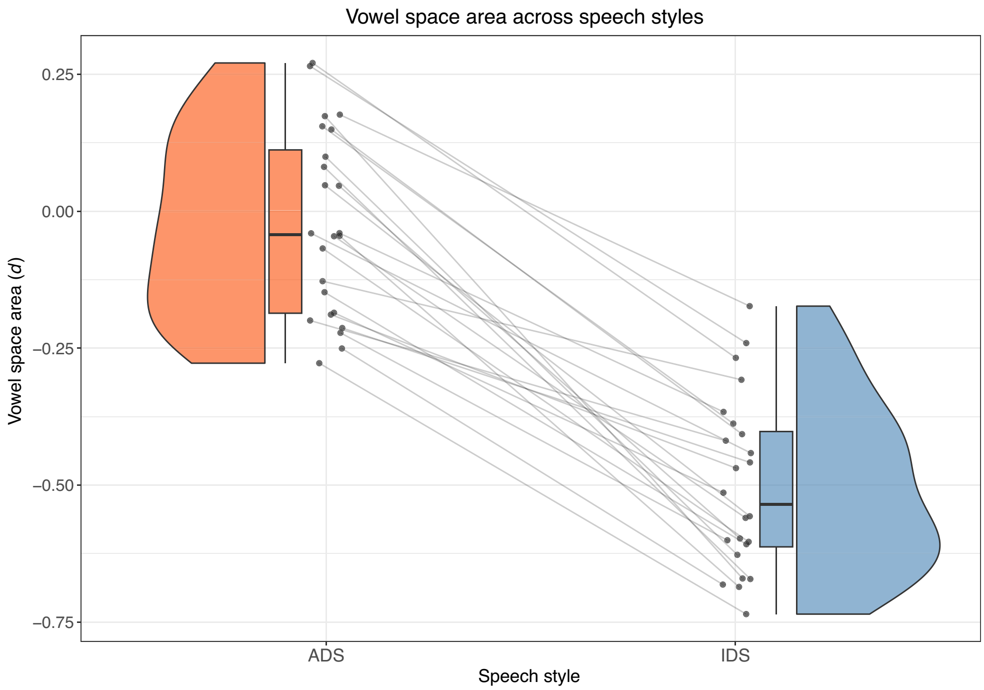

FIGURE 7 Plot of model estimates for subjects' vowel space area across the two speech styles. Each point in each speech style indicates one subject. The points for each subject are connected across the two speech styles with a line. Because model estimates are based on a pooling of repeated measures for each participant, we plot these estimates to provide a more robust picture of differences across speech styles. ADS, adult-directed speech; IDS, infant-directed speech.

area based on all of the border vowels showed no robust differences between the speech styles (estimate = −.14 [−.54; .26], evidence ratio = 2.59; credibility score: 0.72, with 0/26 participants displaying a robust positive effect), nor any effect of infant age (estimate = −.012 [−.046; .022], evidence ratio = 2.59, credibility score: 0.72), as shown in Figures S3.1 and S3.2. We discuss the implications of these results further below. We conducted a control analysis of vowel space area with all border vowels to check whether the speech styles exhibited different values across different contexts of stress, focus, word type, and length (cf. Figure S6.2 for all estimates). Across the majority of contexts, there were no differences in vowel space with all border vowels, with the exception of long vowels where the ADS vowel space area tends to be smaller (estimate = −.44 [−.87; −.01], evidence ratio = 20.2, credibility score = 0.95) as well as unstressed vowels where ADS vowel space area appears to more expanded than in IDS (estimate = .46 [.13; .8], evidence ratio = 75.77, credibility score = 0.99).

Meta-analytic priors

In this section, we consider whether our findings change if we statistically integrate information from a recent meta-analysis of IDS (Cox, Bergmann, et al., 2022) and a prior empirical study of Danish IDS (Bohn, 2013) and how our posterior estimates relate to those different priors. In Figure 8, we depict how each of these priors update into posteriors after seeing the data; that is, we incorporated information from other studies and assessed the extent to which our estimates changed (cf. Fusaroli, Grossman, et al., 2021; Parola et al., 2022 for similar approaches). We chose the parameters of the skeptical prior in this analysis with a view to regularizing the effect of data on the posterior estimates; that is, this skeptical prior encodes the very low likelihood of large effect sizes (cf. Section S1). Figure 8 and Table 4 indicate that prosodic properties of Danish IDS (i.e., f_o, f_o variability and articulation rate) conform to the expectations generated by meta-analytic estimates (in purple), although the estimates for Danish show smaller effect sizes for each of the three acoustic measures (cf. Kvarven et al., 2020). Note, however, that our results for f_o contradict those of Bohn (2013; in green) who found evidence for a null effect. Interestingly, the vowel duration results for Danish contradict the cross-linguistic tendency for vowels to be longer in IDS, as shown by the meta-analytic estimate, but conform to evidence obtained in Bohn's (2013) experimental study of Danish. Lastly, our negative posterior estimate for vowel space area in Danish similarly contradicts cross-linguistic patterns of vowel space expansion, but conforms to the results obtained in Bohn's (2013) experimental study. We should point out that due to the high degree of uncertainty associated with the current sample and the low degree of uncertainty in the meta-analytic estimate based on 32 studies (Cox, Bergmann, et al., 2022), the Danish data do not substantially sway

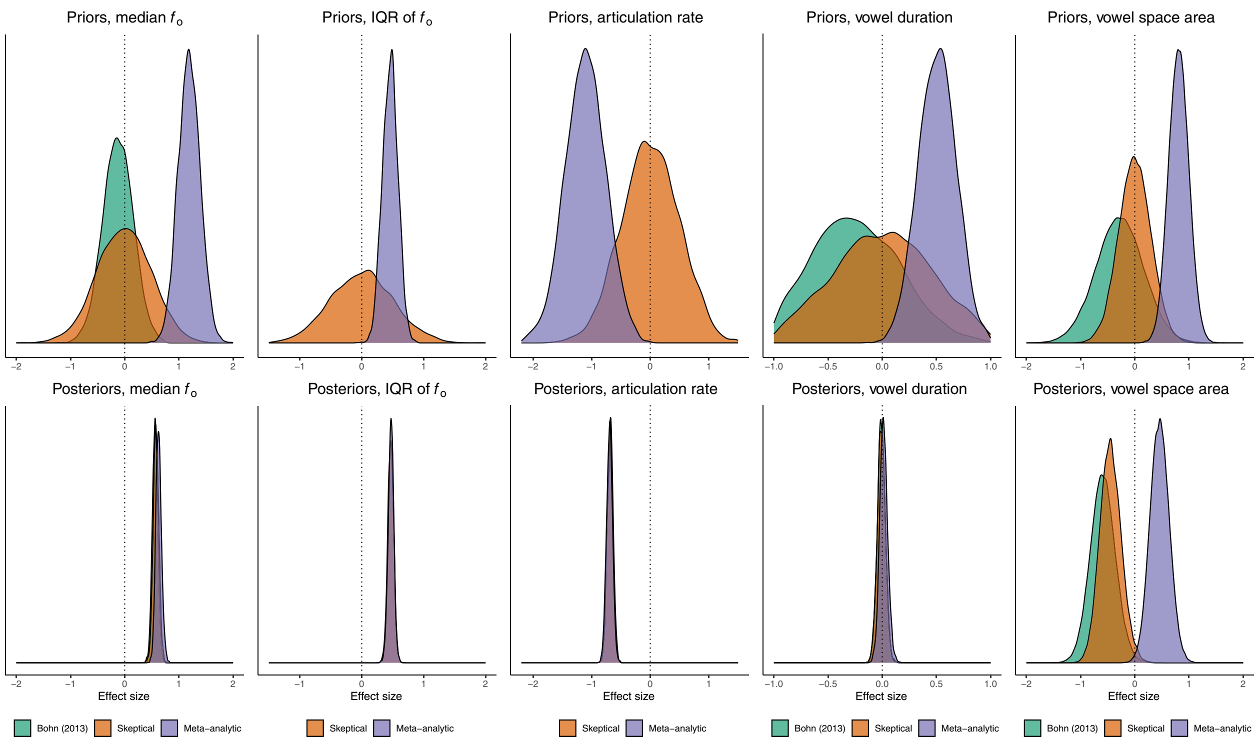

FIGURE 8 A panel of prior–posterior update plots for each of the acoustic measures, showing how the skeptical and meta-analytic priors change as a result of learning from the data. The meta-analytic estimates consist of a synthesis of data from studies on IDS (Cox, Bergmann, et al., 2022) as well as a recent large-scale cross-linguistic study on the features of IDS (Hilton et al., 2021). ADS, adult-directed speech; IDS, infant-directed speech.

the existing evidence of a general vowel space expansion in IDS across languages. The lack of overlap between the skeptical and meta-analytic posterior distributions supports the idea of Danish behaving differently from most other languages in regard to vowel space area in IDS.

Within-vowel variability

This section concerns the third aim of the study, namely to investigate within-vowel variability and between-vowel discriminability in Danish IDS and ADS. The results of the location-scale model pertaining to within-vowel variability indicated that the vowel categories of Danish IDS exhibited a greater degree of within-category variability to those in Danish ADS. This is reflected in Table 5 below, which shows the ratio of evidence in favor of ADS vowel categories showing less variability than IDS vowel categories. The model indicated that all of the vowel clusters in IDS (except /ɑ/) exhibited more variability than in ADS in the dimension of vowel height (i.e., F1), and all of the vowel clusters except /y/ and /ɔ/ exhibit more variability in IDS in the dimension of vowel front-backness (i.e., F2). The model also showed evidence that the majority of the vowels investigated exhibited raising of IDS formants in both height and front-backness dimensions (cf. Table 5). Figure 9 shows posterior samples drawn from the location-scale model of the vowel data and likewise reflects these patterns.

Vowel discriminability measure

Similarly, the model of vowel discriminability indicates that caregivers produce less discriminable vowels in IDS (0.54 [0.50; 0.59]) than in ADS (0.67 [0.64; 0.70]), as shown in Figure 10. We obtain strong evidence in favor of the hypothesis that caregivers exhibit a lower proportion of explained variance in IDS (−0.13 [−0.16; −0.09], evidence ratio = Inf, credibility score = 1, with 24/26 participants displaying a reliable negative effect). The model also indicates a similar degree of variability in the distribution of values in IDS (σ = 0.023 [0.002; 0.088]) and in ADS (σ = 0.017 [0.002; 0.062]). We accordingly obtain weak evidence in favor of the hypothesis that IDS exhibits a greater amount of variability (0.01 [−0.04; 0.06], evidence ratio = 0.7, credibility score: 0.7, with 0/26 participants showing a reliable positive effect). The model with age as a predictor exhibits similar out-of-sample predictive accuracy (elpd: −1.2, se: 4.7) compared to the model without age as a predictor; however, as shown in Figure S2, we see no robust effect of age (estimate = .001 [−.007; .009], evidence ratio = 0.63, credibility score = 0.38).

TABLE 4 Effect size estimates for each of the acoustic features in Danish infant-directed speech (IDS). The columns compare the posterior estimates for the models with meta-analytic priors (Cox, Bergmann, et al., 2022), priors from Bohn's (2013) study, and the models with skeptical priors (cf. S1.1 in Supporting Information for more information about the skeptical priors used here). The meta-analytic estimates consist of a synthesis of data from studies on IDS (Cox, Bergmann, et al., 2022) as well as a recent large-scale cross-linguistic study on the features of IDS (Hilton et al., 2021).

Acoustic measure	Estimates from Bohn (2013)	Meta-analytic estimates from Cox, Bergmann, et al. (2022)	Posterior estimate (w. meta-analytic prior)	Posterior estimate (w. Bohn (2013) prior)	Posterior estimate (w. skeptical prior)
Median f_o	0.13 [0.41; 0.67]	1.20 [0.78; 1.62]	0.63 [0.53; 0.74]	0.56 [0.45; 0.66]	0.58 [0.47; 0.68]
IQR of f_o	NA	0.49 [0.15; 0.82]	0.48 [0.38; 0.57]	NA	0.47 [0.37; 0.56]
Articulation rate	NA	−1.12 [−1.76; −0.42]	−0.72 [−0.82; −0.61]	NA	−0.68 [−0.78; −0.57]
Vowel duration	0.20 [0.34; 0.74]	0.52 [0.22; 0.83]	−0.02 [−0.05; 0.10]	−0.01 [−0.08, 0.07]	−0.01 [−0.08; 0.06]
Vowel space area	−0.28 [−1.14; 0.58]	0.83 [0.53; 1.13]	0.59 [0.30; 0.87]	−0.58 [−1.02; −0.12]	−0.39 [−0.77; 0.03]

DISCUSSION

This paper set out to investigate caregivers' spontaneous IDS and ADS in Danish and quantify the extent to which caregivers modify the prosodic and vocalic features across the two speech styles. We turned our focus to three different aspects of the acoustic expression of Danish IDS. First, we looked at the extent to which five acoustic properties that have been reported extensively in the literature on IDS (Cox, Bergmann, et al., 2022) were expressed in Danish. The results indicated that Danish caregivers conformed to cross-linguistic patterns of prosodic properties of IDS, producing IDS with a higher median f_o, a greater degree of f_o variability, and a slower rate of articulation. The vocalic measures of vowel duration and vowel space area in IDS, on the other hand, contradicted cross-linguistic tendencies, with caregivers producing either a reduced or similar vowel space area in IDS (cf. Figure 7 and Figure S3.1), as well as similar vowel durations or slightly longer vowels in ADS (cf. Figure 6; Figure S6.1). Second, we asked whether any of the acoustic properties of Danish caregivers' IDS exhibited change according to infant age. The results indicated no clear age-related changes in the majority of the prosodic and vocalic measures (cf. S2 in Supporting Information), with the exception of articulation rate, which became gradually more similar to Danish ADS in IDS addressed to older infants (cf. Figure 5). Lastly, we asked how parents negotiated the balance between within-vowel variability and between-vowel separability, and whether vowel categories in Danish IDS exhibited a higher degree of discriminability. In our sample of participants, Danish caregivers produced formant raising, a greater degree of within-vowel variability and a lower degree of between-vowel discriminability in IDS when compared with ADS. While our results thus partially support prior cross-linguistic findings (e.g., Englund & Behne, 2005; Rattanasone et al., 2013; Rosslund et al., 2022), they also provide further evidence of the unusual nature of Danish speech and conversational practices (cf. Dideriksen et al., 2022; Trecca et al., 2021 for a review). In the following discussion, we discuss each of these results in turn and argue that they call for more hypothesis-driven comparisons between similar languages (Christiansen et al., 2022) as well as the widespread adoption of cumulative science practices (Brand et al., 2019; Fusaroli, Grossman, et al., 2021).

Danish IDS and cross-linguistic patterns

The finding that Danish caregivers produce speech with an elevated pitch, more varied melody and fewer syllables per second in IDS than in ADS conforms to cross-linguistic patterns of prosodic features of IDS (Cox, Bergmann, et al., 2022; Hilton et al., 2022). The

TABLE 5 Evidence ratios for individual vowel categories. An evidence ratio of >1 indicates evidence in favor of the hypothesis either (i) that the value of F1 or F2 in ADS is lower than that in IDS (columns 2 and 3) or (ii) that the F1 or F2 of the vowels in ADS exhibits less variability than that in IDS (columns 4 and 5).

Vowel cluster	$F1_{ADS} < F1_{IDS}$	$F2_{IDS} > F2_{ADS}$	F1, $\sigma_{ADS} < \sigma_{IDS}$	F2, $\sigma_{ADS} < \sigma_{IDS}$
i	189	0.38	Inf	399
y	31.5	9.84	7.97	0.77
e	Inf	1.76	Inf	132.00
ø	3999	443.00	3999	9.05
ɛ	Inf	3.19	Inf	25.0
æ	20.86	0.906	3.23	1.53
œ	5.68	Inf	0.922	38.2
ɑ	1.56	Inf	56.1	147.00
ʌ	Inf	Inf	Inf	69.2
ɔ	Inf	799	332	0.269
o	49.6	147.00	1.05	4.47
u	60.5	29.3	18.0	147.0

Abbreviations: ADS, adult-directed speech; IDS, infant-directed speech.

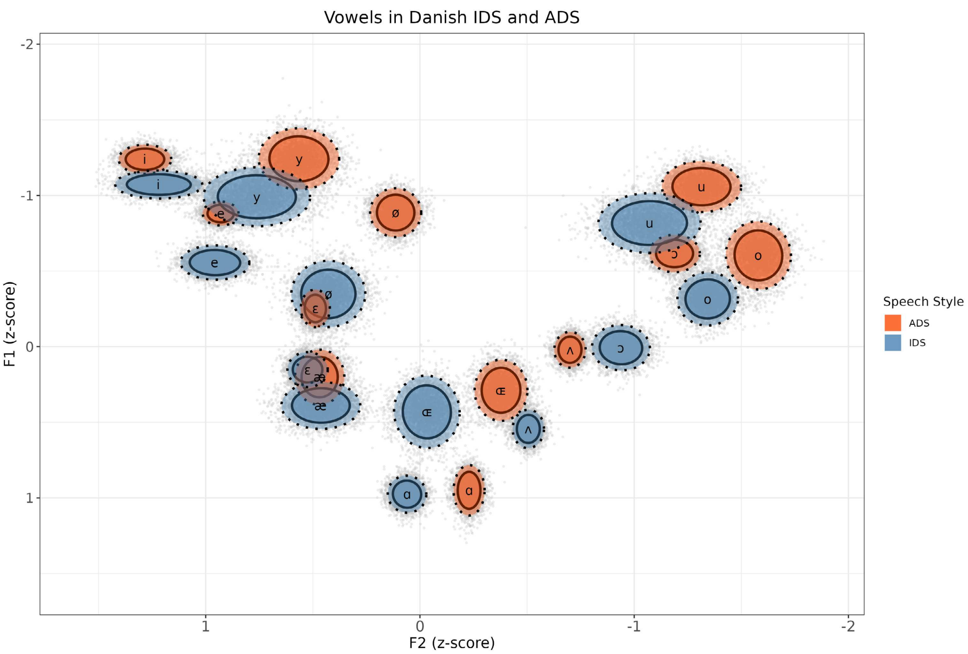

FIGURE 9 Plot of posterior draws from the location-scale model of vowel categories in Danish IDS (blue) and ADS (orange). The dotted ellipsis encompasses 95% of the vowels while the innermost ellipsis surrounds 80% of the vowel distribution. Note that we only model vowel categories for which we have above 50 tokens (cf. Table 3). ADS, adult-directed speech; IDS, infant-directed speech.

current results provided further evidence showing that the speech style we use when interacting with young children has similar prosodic features across a wide intersection of languages (Broesch & Bryant, 2015; Bryant & Barrett, 2007; Hilton et al., 2022). The above results for Danish and most cross-linguistic studies highlight pitch properties as being the most salient cues in IDS during early development (Broesch & Bryant, 2015; Bryant & Barrett, 2007; Fernald, 1989). These acoustic properties serve the functions of communicating intentions, grabbing an infants' attention, expressing emotions, and encouraging behavior (Bryant & Barrett, 2007; Fernald & Mazzie, 1991; Fernald & Simon, 1984). This result also reverberates with studies showing that the tendency for IDS to grab infants' attention seems to be driven primarily by pitch elevation (Segal & Newman,

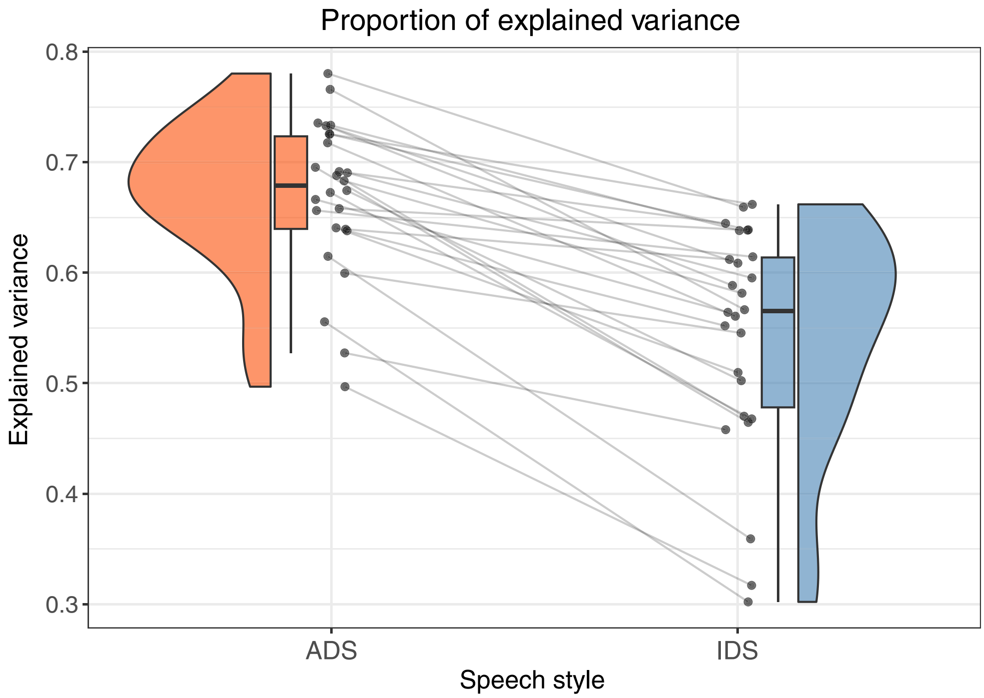

FIGURE 10 Plot of model estimates for the proportion of explained variance from category membership across the two speech styles. Each point in each speech style indicates one subject. The points for each subject are connected across the two speech styles with a colored line. ADS, adult-directed speech; IDS, infant-directed speech.

2015). The finding of a higher median f_o in this study actually contradicts a previous study of Danish IDS (Bohn, 2013); however, the cause of this discrepancy may be due to methodological differences, as we know that recordings of spontaneous speech produces bigger effect sizes compared to those of more controlled speech (Cox, Bergmann, et al., 2022). The explicit integration of Bohn's (2013) results into the prior of our statistical model allowed us to quantify the extent to which this prior updates into a positive posterior estimate after seeing our data (cf. Figure 8). Given the small differences between the posterior estimates with priors from Cox, Bergmann, et al.'s (2022) meta-analysis and Bohn's (2013) experimental study, the results suggest that we now have strong evidence that f_o in Danish IDS is higher than in ADS. It should be noted, however, that the size of the effect is more moderate than the effect size suggested by the meta-analysis of IDS (Cox, Bergmann, et al., 2022).

The greatest difference between the two speech styles was a slower articulation rate in IDS (cf. Table 4). This acoustic property of Danish IDS again conforms to cross-linguistic patterns of acoustic features of IDS (Cox, Bergmann, et al., 2022) and may serve the purpose of easing the cognitive demand involved in young infants' processing of speech (Christiansen & Chater, 2016; Peter et al., 2016; Saffran & Kirkham, 2018; Thiessen et al., 2005). A slowed articulation rate has also been shown to increase the intelligibility of speech (Ferguson & Kewley-Port, 2007; Lam & Tjaden, 2013; Searl & Evitts, 2013), to facilitate word recognition and learning (Raneri, 2015; Song et al., 2010) and to be used when introducing unfamiliar words (Han et al., 2018). In terms of prosodic properties of IDS, then, Danish caregivers made acoustic modifications in a way that suggests flexible adaptation to infants' communicative immaturity and developmental needs (Fusaroli, Weed, et al., 2021; Goldstein & Schwade, 2008; Ko et al., 2016; Nguyen et al., 2022; Warlaumont et al., 2014).

Our analysis of vowel space area in Danish IDS and ADS contradicted the general cross-linguistic tendency for vowel space expansion in IDS (e.g., Hartman et al., 2017; Liu et al., 2003), but was consistent with earlier acoustic investigations of IDS in Danish (Bohn, 2013; Dideriksen & Fusaroli, 2018) and the growing number of studies showing language specificity in some of the acoustic properties of IDS (e.g., Englund, 2018; Rattanasone et al., 2013; Rosslund et al., 2022). Vowel space reduction, for example, has been found in a number of different languages, such as Dutch (Benders, 2013), Norwegian (Englund, 2018; Englund & Behne, 2005; Steen & Englund, 2021), and Cantonese (Rattanasone et al., 2013). Considering the general connection between speech clarity and vowel space expansion (Lam et al., 2012; Whitfield & Goberman, 2017; Whitfield & Mehta, 2019)—as well as the phonetic opacity of Danish sound structure—we may have expected Danish caregivers to increase the separability between the centroids of vowel categories in IDS. Our control analysis of the vowel space area measures (cf. Figure S6.2) shows that even in contexts of emphasis (e.g., focused constituents and content words), we see no clear evidence of vowel space expansion, the only potential exception being long vowel contexts where the vowel space in IDS is slightly expanded. These results highlight the need for more studies to test whether clear speech in Danish as well as other languages can be described with vowel space expansion, or to a greater extent can be measured by a slowed articulation rate or some other mediating acoustic variable (Ferguson & Kewley-Port, 2007; Lam & Tjaden, 2013; Searl & Evitts, 2013).

Another feature that has been proposed to aid speech intelligibility involves an exaggeration of differences in vowel duration, thus making relevant phonological differences more salient to children (Seidl & Cristià, 2008; Soderstrom et al., 2003). In our first control analysis of the vowels across different contexts of phonological length (cf. Figure S6.1), we found that phonologically long vowels in ADS exhibit longer duration compared to those in IDS, whereas vowels of short phonological length exhibited similar durations across the two speech styles. These findings of no clear differences between vowel durations across the two speech styles should be interpreted in light of the use of quantity distinctions to distinguish word meaning in Danish. For example, distinctions between the singular *tov* [tɒʊ̯] “rope” and the plural *tove* [ˈtɒʊ̯ː] “ropes” or between *hus* [huːˀs] “house” and *huse* [huːːs] “houses” often relies on subtle quantity differences in vowel length. The finding of a lack of exaggeration of IDS vowel lengths within this complex quantity system may suggest that caregivers rely on more transparent cues to clarify these word meanings (Kjærbæk & Basbøll, 2016), such as clearer consonantal cues or [ɐ]- and [ə]-suffixation (e.g., *tove* [ˈtʌwə] “ropes” and *huse* [ˈhuːsə] “houses”). Moreover, because longer segmental duration tends to co-occur with a slowed articulation rate (Panneton et al., 2006; Song et al., 2010), we conducted a second control analysis where we normalized vowel duration by the inverse of the median articulation rate for each speaker. Rate-normalized vowel length was shorter in IDS than in ADS (cf. Figure S9.1), implying that vowels constitute a smaller proportion of the speech stream in IDS. Because rate-normalized vowel length exhibited similar slopes for phonologically long and short vowels across speech styles (cf. Figure S9.2), the results still suggest a lack of exaggerated quantity distinctions between long and short vowels in IDS. These findings hold interesting implications in the context of Danish phonetic structure, where the frequent reduction of obstruents to vocalic sounds (i.e., consonant reduction) has been argued to reduce the salience of cues that allow infants to extract information from the speech stream (Bleses et al., 2008a, 2008b; Højen & Nazzi, 2016; Trecca et al., 2019). The finding that rate-normalized vocalic material comprises a smaller part of the speech stream in IDS may tentatively suggest that Danish caregivers to a greater extent exaggerate consonantal cues in IDS, presumably to aid Danish-learning infants in segmenting a highly vocalic speech stream (Bleses et al., 2008a; Trecca et al., 2019, 2020). These intriguing findings in Danish IDS and ADS vowels require further experimental investigation and demonstrate the importance of considering subtle distinctions across phonological systems in discovering differences in the acoustic expression of IDS and ADS.

Stability and dynamic changes in features

The second aim of this paper was to investigate the extent to which the acoustic expression of Danish IDS changes with infant age. The results indicated no clear age-related changes in the majority of the prosodic and vocalic measures (cf. S2 in Supporting Information), with the exception of articulation rate, which became gradually more similar to Danish ADS the older the infants were (cf. Figure 5). This finding of an age-related change conformed to other cross-linguistic studies of age-related changes in the acoustic properties of IDS (Kondaurova et al., 2013; Lee et al., 2014; Narayan & McDermott, 2016; Raneri, 2015) and may reflect caregivers' adaptation to infants' gradual improvement in their processing of the speech stream over the span of early development (Christiansen & Chater, 2016; Peter et al., 2016; Saffran & Kirkham, 2018; Werker & Tees, 1999). The lack of change in vowel duration, however, which often goes hand in hand with articulation rate, contradicted longitudinal studies in several languages, indicating that caregivers often decreased the relative vowel duration differences in IDS and ADS as infants became older (Englund & Behne, 2005; Hartman et al., 2017; Vosoughi & Roy, 2012). The lack of age-related changes in pitch properties likewise contradicted the majority of studies that suggest that caregivers reduce the median f_o and f_o variability in IDS as their infants become older (Amano et al., 2006; Gergely et al., 2017; Han et al., 2020; Kondaurova et al., 2013; Niwano & Sugai, 2002; Stern et al., 1983; Vosoughi & Roy, 2012). Similarly, we saw no changes in either vowel space expansion or vowel discriminability. The stability in these measures may indicate their continued function of increasing infant attention and social motivation as well as expressing affect (Fernald, 1989; Fernald & Kuhl, 1987; Kitamura & Lam, 2009; although see Ma et al., 2011); however, any conclusions that we can make about age-related changes may be limited by our reliance on a cross-sectional sample, as discussed further in the *Limitations & Future Directions* section below.

Within-vowel variability and between-vowel discriminability

One of the main objectives of this study was to broaden our knowledge of the internal distributions of vowel categories in Danish IDS and ADS by investigating each of the members in the large vowel inventory. The results indicated that caregivers produce more variable vowel categories in IDS compared to ADS, providing another example to add to the growing number of studies showing less compact vowel categories in IDS (cf. Cristia & Seidl, 2014; Martin et al., 2015; McMurray et al., 2013; Miyazawa et al., 2017; Rosslund et al., 2022). This larger degree of variability in IDS vowel categories also influenced the between-vowel discriminability in IDS, with vowels being more overlapping in IDS compared to ADS. This lower degree of vowel discriminability did not appear to change across infant ages (cf. Figure S2).

The combination of less between-vowel separability (i.e., a similar or reduced vowel space area) and a higher degree of within-vowel variability in IDS at first glance contradicted the prominent hypothesis that IDS serves to clarify the speech signal and help infants learn phonetic categories (Hartman et al., 2017; Liu et al., 2003). However, the presence of variability may benefit infants in a number of different ways. One way that variability may benefit infants is by leading the infant to a greater degree of abstraction from individual categories and a more robust system of categorization (Perry et al., 2010; Raviv et al., 2022; Rost & McMurray, 2009, 2010). For example, Rost and McMurray (2009) found that infants trained on labels spoken by a single speaker failed to distinguish between labels for visual objects that had a minimal phonological difference, whereas infants who were exposed to labels uttered by multiple speakers succeeded. Houston and Jusczyk (2000) similarly showed that increasing the number of speakers during familiarization facilitated 5- to 7-month-old infants' generalization of sound patterns to novel speakers. Infant word learning has likewise been shown to rely on variability; Perry et al.'s (2010) longitudinal study showed that infants trained on a more variable set of stimuli exhibited greater generalization to novel stimuli of these categories. These studies indicate that a greater degree of variability may allow infants to abstract away from instances that are not good exemplars of category (Eaves et al., 2016).

The above notions resonate with theories suggesting that infants identify phonological distinctions by observing and processing statistical regularities in the speech stream (e.g., Pierrehumbert, 2003). This construal of language development posits that infants form a phonological system through a gradual process of matching acoustic input with memories of similar events in an active constructive process. Infants' perceptual development may therefore admit a crucial role for variability, as within-category variability can allow infants to compute statistics over multiple dimensions concurrently (Pierrehumbert, 2003).

Vowel variability may benefit infant learning in a second way: Lower levels of variability and complexity can lead to habituation, which can counteract learning by causing low attention (Colombo & Mitchell, 2009; Hunter et al., 1983; Paulus, 2022). In this sense, a greater degree of variability may be necessary for attention and learning in phonetic category development (Christiansen & Chater, 2016; Raviv et al., 2022). For example, an experimental study has shown that 14-month-old infants learned faster from a single speaker who spoke with a greater degree of variability in duration, pitch, and pitch variability (Galle et al., 2015). The notion that variability in the speech stream can attract infants' attention goes hand in hand with the results obtained in this study that f_o, f_o variability and articulation rate all exhibited a greater degree of variability within Danish IDS than within ADS. The beneficial role of variability may thus consist primarily in its ability to grab and maintain infant attention (Englund, 2018).

A third possibility is that the greater degree of vowel variability could potentially be a side effect of other articulatory features specific to IDS. This may include an elevated pitch and pitch variability, which has been shown to impact both F1 and F2 measures (McMurray et al., 2013). Relatedly, the tendency for caregivers to raise their larynx in IDS—either to convey non-threatening behavior (Kalashnikova et al., 2017), to mimic infant production (Cristia, 2013; Polka et al., 2022), to grab infant attention (Masapollo et al., 2016), or to convey positive affect (Benders, 2013; Saint-Georges et al., 2013; Singh et al., 2002), or a combination thereof—would produce a shorter vocal tract and result in an increase in both the first and second formants for all of the vowels. The results indeed indicated a leftward and downward transformation (i.e., raised F1 and F2 values) in the IDS formant centroids for the majority of the vowel categories investigated (cf. Table 5).

Another articulatory factor that might shorten the vocal tract and produce the shifts in formant frequencies would be through smiling. Smiling involves a retraction of the lips and widening of the mouth, with a resultant increase in the first and second vowel formants (Barthel & Quené, 2015; Tartter, 1980). The acoustic origin of smiling in the animal kingdom has been posited to derive from the desire to raise the resonant frequencies of the vocal tract to sound smaller and convey appeasement toward others (Ohala, 1980; Xu & Chuenwattanapranithi, 2007). This last explanation would admit a crucial role for the audio–visual component of language development and would conform to evidence indicating that infants can integrate audio–visual speech stimuli at an early point in development (Cox, Keren-Portnoy, et al., 2022) and flexibly take advantage of intersensory redundancy when processing complex audio–visual speech stimuli (Bastianello et al., 2022; Hillairet de Boisferon et al., 2017; Lewkowicz & Hansen-Tift, 2012; Pons et al., 2015). Infants also preferentially attend to infant-directed faces (Kim & Johnson, 2014), and smiling during interaction produces a greater frequency of speechlike syllabic infant vocalizations (Hsu et al., 2001). The observed formant raising in IDS could thus be in part motivated by these visual and emotional accompaniments to verbal communication (Englund, 2018). It is important to recognize that raising the formants in IDS does not necessarily mean that caregivers cannot simultaneously expand the vowel space. These two features of IDS can work together, and exploring how they interact and contribute to the proposed functions of IDS will be a valuable topic for future research.

Limitations and future directions

The results of this study highlight the need for detailed cross-linguistic analysis of the acoustic properties of IDS and demonstrate the value of comparatively and critically incorporating statistical information from prior studies (Brand et al., 2019; Devezer et al., 2019; Fusaroli, Grossman, et al., 2021). There are several limitations of this study, which are important to keep in mind for planning future studies. The first limitation concerns the characteristics of our participant sample. The lack of age-related effects for most of the acoustic features of IDS examined here must be interpreted in light of relying on a cross-sectional sample, as some of the ages were represented by only one or two participants. Because the infants of the caregivers under investigation were between 11 and 24 months of age, moreover, they would already have a certain level of exposure and knowledge of the phonetic categories of their first language (cf. Kuhl, 2000). The current results, for example, cannot rule out the possibility that caregivers initially produce a greater degree of vowel separability to younger infants, as indicated in studies of other languages (cf. Hartman et al., 2017; Kuhl et al., 1997; H. Liu et al., 2003). If IDS is construed as a form of speaker adaptation, we might also expect infants' developmental status as well as the kinship status and familiarity of the interlocutor to affect the generalizability of the results. Future research exploring the effects of diverse speaker characteristics would provide important insights into factors affecting the acoustic properties of IDS (e.g., Kaplan et al., 2001; Lam-Cassettari & Kohlhoff, 2020; Steen & Englund, 2021; Woolard et al., 2022). The recent expansion in the availability of cross-linguistic data (e.g., MacWhinney, 2014) and technological improvements to perform automatic transcription (e.g., Cychosz et al., 2021; Räsänen et al., 2021) allow more detailed analyses of how IDS differs across individuals, genders, languages, and infant ages. To take full advantage of cumulative science practices, we would encourage researchers to share utterance- and vowel-level data in open repositories. This approach to future investigations of IDS would allow for a higher resolution of how the acoustic properties of IDS differ between individuals, languages, and infant ages. In line with this recommendation, the data and code used in this manuscript are available in the following open repository: https://osf.io/ywf9m/?view_only=d99fc6dbc61546febff619b8674a7943.

A second limitation of this study concerns the focus on Danish without any comparison to other languages or cultures that differ in a key moderator of interest. For example, a comparison between Danish and a similar vowel-rich language without quantity distinctions would provide important insights into how phonological structure can influence the acoustic expression of IDS. By conducting theory-driven comparisons of the acoustic properties of IDS across a diverse intersection of languages and cultures (Christiansen et al., 2022; Deffner et al., 2021), we can obtain a fuller picture of the cross-cultural and cross-linguistic variables that moderate the acoustic expression of the speech style. One approach would be to investigate culturally similar societies that differ in the key phonological variables of interest (e.g., number of front vowels, consonantal lenition, or schwa assimilation). By keeping unobserved cultural variables roughly comparable, such as socio-economic status and child-rearing practices, we can isolate the influence of language structure and facilitate causal inference about specific acoustic modifications in caregivers' IDS. Another approach would include computational models with IDS and ADS speech data from a diverse intersection of languages to formulate testable predictions on whether the benefits of the speech style derive from the improved information structure of the speech signal (e.g., Eaves et al., 2016; Ludusan et al., 2021; McMurray et al., 2009). This in turn would allow us to examine and answer questions as to whether caregivers' acoustic modifications in IDS relate mainly to phonological structure or cultural practice.

CONCLUSION

The present study was designed as a comparative analysis of the acoustic properties of Danish ADS and IDS. First, the results indicated that pitch, melody, and articulation rate in IDS were modified by Danish caregivers in similar ways to other languages. However, Danish vowels in IDS were articulated with no acoustic exaggeration and with similar durations compared to ADS. The findings for the vocalic properties of Danish IDS here thus add to a small subset of studies finding no vocalic exaggeration in IDS and provide further evidence of the peculiar nature of Danish sound structure. Second, articulation rate was the only property of Danish IDS to exhibit dynamic change and became more similar to Danish ADS when directed to older infants. Third, Danish caregivers produced a greater degree of within-vowel variability and a lower degree of between-vowel discriminability in IDS when compared with ADS. These findings highlight the need for future studies to conduct theory-driven comparisons of the acoustic expression of IDS across a wide intersection of languages with distinct phonological systems.

ACKNOWLEDGMENTS

This research was supported in part by the Danish Council for Independent Research (FKK Grant

DFF-7013-00074) to MHC. This project also received a seed funding grant from the Interacting Minds Centre. All of the computation done for this project was performed on the UCloud interactive HPC system, which is managed by the eScience Center at the University of Southern Denmark. We also thank Ocke-Schwen Bohn for his insightful comments on an earlier version of this paper.

DATA AVAILABILITY STATEMENT

The data necessary to reproduce the analyses presented here are publicly accessible at the following URL: https://osf.io/ywf9m/?view_only=d99fc6dbc61546febff619b8674a7943. The analytic code necessary to reproduce the analyses presented in this paper is publicly accessible at: https://osf.io/ywf9m/?view_only=d99fc6dbc61546febff619b8674a7943. The materials (i.e., videos and audio files) necessary to attempt to replicate the findings presented here are not publicly accessible. The analyses presented here were not preregistered.

ORCID

Christopher Cox https://orcid.org/0000-0003-4076-581X
Christina Dideriksen https://orcid.org/0000-0002-0528-5530
Tamar Keren-Portnoy https://orcid.org/0000-0002-7258-2404
Andreas Roepstorff https://orcid.org/0000-0002-3665-1708
Morten H. Christiansen https://orcid.org/0000-0002-3850-0655
Riccardo Fusaroli https://orcid.org/0000-0003-4775-5219

REFERENCES

Adank, P., Smits, R., & Van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, *116*(5), 3099–3107.

Amano, S., Nakatani, T., & Kondo, T. (2006). Fundamental frequency of infants' and parents' utterances in longitudinal recordings. *The Journal of the Acoustical Society of America*, *119*(3), 1636–1647. https://doi.org/10.1121/1.2161443

Barthel, H., & Quené, H. (2015). Acoustic-phonetic properties of smiling revised: Measurements on a natural video corpus. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences* (pp. 1–5). The University of Glasgow.

Basbøll, H. (2005). *The phonology of Danish*. Oxford University Press.

Bastianello, T., Keren-Portnoy, T., Majorano, M., & Vihman, M. (2022). Infant looking preferences towards dynamic faces: A systematic review. *Infant Behavior and Development*, *67*(101), 709.

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

Bergelson, E., 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*(1), e12724.

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences.*, *26*, 1153–1170. https://doi.org/10.1016/j.tics.2022.09.015

Bleses, D., Basbøll, H., & Vach, W. (2011). Is Danish difficult to acquire? Evidence from Nordic past-tense studies. *Language and Cognitive Processes*, *26*(8), 1193–1231.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008a). Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of Child Language*, *35*(3), 619–650.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008b). The Danish communicative developmental inventories: Validity and main developmental trends. *Journal of Child Language*, *35*(3), 651–669.

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* [computer program]. Version 6.2.05. http://www.praat.org/

Bohn, O.-S. (2013). Acoustic characteristics of Danish infant directed speech. *Proceedings of Meetings on Acoustics ICA2013* (Vol. 19, no. 1, p. 060055), Acoustical Society of America.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities. *Journal of Speech, Language, and Hearing Research*, *46*(1), 80–97.

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, *20*(3–4), 255–272.

Brand, C. O., Ounsley, J. P., van der Post, D. J., & Morgan, T. J. H. (2019). Cumulative science via Bayesian posterior passing: An introduction. *Meta-Psychology*, *3*. https://doi.org/10.15626/MP.2017.840

Broesch, T. L., & Bryant, G. A. (2015). Prosody in infant-directed speech is similar across western and traditional cultures. *Journal of Cognition and Development*, *16*(1), 31–43.

Bryant, G. A. (2022). Vocal communication across cultures: Theoretical and methodological issues. *Philosophical Transactions of the Royal Society B*, *377*(1841), 20200387. https://doi.org/10.1098/rstb.2020.0387

Bryant, G. A., & Barrett, H. C. (2007). Recognizing intentions in infant-directed speech: Evidence for universals. *Psychological Science*, *18*(8), 746–751. https://doi.org/10.1111/j.1467-9280.2007.01970.x

Burnham, E. B., Wieland, E. A., Kondaurova, M. V., McAuley, J. D., Bergeson, T. R., & Dilley, L. C. (2015). Phonetic modification of vowel space in storybook speech to infants up to 2 years of age. *Journal of Speech, Language, and Hearing Research*, *58*(2), 241–253.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Tseltal Mayan village. *Child Development*, *91*(5), 1819–1835. https://doi.org/10.1111/cdev.13349

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, E62. https://doi.org/10.1017/S0140525X1500031X

Christiansen, M. H., Kallens, P. C., & Trecca, F. (2022). Towards a comparative approach to language acquisition. *Current Directions in Psychological Science*, *31*, 131–138. https://doi.org/10.1177/09637214211049229

Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory*, *92*(2), 225–234.

Cox, C. M. M., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G. A., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour*, *7*, 114–133. https://doi.org/10.1038/s41562-022-01452-1

Cox, C. M. M., Keren-Portnoy, T., Roepstorff, A., & Fusaroli, R. (2022). A Bayesian meta-analysis of infants' ability to perceive audio–visual congruence for speech. *Infancy*, *27*(1), 67–96.

Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass*, *7*(3), 157–170.

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

Cychosz, M., Cristia, A., Bergelson, E., Casillas, M., Baudet, G., Warlaumont, A. S., Scaff, C., Yankowitz, L., & Seidl, A. (2021). Vocal development in a large-scale crosslinguistic corpus. *Developmental Science*, *24*, e13090. https://doi.org/10.1111/desc.13090

De Boer, B., & Kuhl, P. K. (2003). Investigating the role of infant-directed speech with a computer model. *Acoustics Research Letters Online*, *4*(4), 129–134. https://doi.org/10.1121/1.1613311

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385–390.

Deffner, D., Rohrer, J. M., & McElreath, R. (2021). A causal framework for cross-cultural generalizability. *Advances in Methods and Practices in Psychological Science*, *5*(3), 25152459221106366. https://doi.org/10.1177/25152459221106366

Degottex, G., Kane, J., Drugman, T., Raitio, T., & Scherer, S. (2014). *COVAREP—A collaborative voice analysis repository for speech technologies*. IEEE. https://doi.org/10.1109/ICASSP.2014.6853739

Devezer, B., Nardin, L. G., Baumgaertner, B., & Buzbas, E. O. (2019). Scientific discovery in a model-centric framework: Reproducibility, innovation, and epistemic diversity. *PLoS One*, *14*(5), e0216125.

Dideriksen, C., Christiansen, M. H., Dingemanse, M., Højmark-Bertelsen, M., Johansson, C., Tylén, K., & Fusaroli, R. (2022). Language specific constraints on conversation: Evidence from Danish and Norwegian. *PsyArxiv*.

Dideriksen, C., & Fusaroli, R. (2018). *CDS is not what you think-hypoarticulation in Danish child directed speech*. Poster presented at Nijmegen Lectures.

Dodane, C., & Al-Tamimi, J. (2007). An acoustic comparison of vowel systems in adult-directed-speech and child-directed-speech: Evidence from French, English & Japanese. *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, Newcastle University.

Eaves, B. S., Feldman, N. H., Griffiths, T. L., & Shafto, P. (2016). Infant-directed speech is consistent with teaching. *Psychological Review*, *123*(6), 758–771. https://doi.org/10.1037/rev0000031

Englund, K. T. (2018). Hypoarticulation in infant-directed speech. *Applied Psycholinguistics*, *39*(1), 67–87.

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

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *112*(1), 259–271.

Ferguson, S. H., & Kewley-Port, D. (2007). Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research*, *50*(5), 1241–1255.

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

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, *10*(3), 279–293.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, *27*(2), 209–221.

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

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

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

Fusaroli, R., Grossman, R., Bilenberg, N., Cantio, C., Jepsen, J. R. M., & Weed, E. (2021). Toward a cumulative science of vocal markers of autism: A cross-linguistic meta-analysis-based investigation of acoustic markers in American and Danish autistic children. *Autism Research*, *15*, 653–664. https://doi.org/10.1002/aur.2661

Fusaroli, R., Weed, E., Fein, D., & Naigles, L. (2019). Hearing me hearing you: Reciprocal effects between child and parent language in autism and typical development. *Cognition*, *183*, 1–18. https://doi.org/10.1016/j.cognition.2018.10.022

Fusaroli, R., Weed, E., Rocca, R., Fein, D. & Naigles, L. (2023). Caregiver linguistic alignment to autistic and typically developing children: A natural language processing approach illuminates the interactive components of language development. *Cognition*, *236*, 105422.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *182*(2), 389–402.

Galle, M. E., Apfelbaum, K. S., & McMurray, B. (2015). The role of single talker acoustic variation in early word learning. *Language Learning and Development*, *11*(1), 66–79.

García-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology*, *110*, 1–17.

García-Sierra, A., Ramírez-Esparza, N., Wig, N., & Robertson, D. (2021). Language learning as a function of infant directed speech (IDS) in Spanish: Testing neural commitment using the positive-MMR. *Brain and Language*, *212*(104), 890.

Gelman, A., Simpson, D., & Betancourt, M. (2017). The prior can often only be understood in the context of the likelihood. *Entropy*, *19*(10), 555.

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. https://doi.org/10.1038/s41598-017-13883-2

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

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

Grønnum, N. (1998). Danish. *Journal of the International Phonetic Association*, *28*(1–2), 99–105.

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

Han, M., de Jong, N. H., Kager, R., & Bertolini, A. B. (2018). Infant-directed speech is not always slower: Cross-linguistic evidence from Dutch and Mandarin Chinese. *Proceedings of the 42nd Annual Boston University Conference on Language Development*, *2*(1), 331–344.

Hartman, K. M., Ratner, N. B., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language*, *44*(5), 1140–1162. https://doi.org/10.1017/S0305000916000520

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics*, *64*(2), 627–634.

Hillairet de Boisferon, A., Tift, A. H., Minar, N. J., & Lewkowicz, D. J. (2017). Selective attention to a talker's mouth in infancy: Role of audiovisual temporal synchrony and linguistic experience. *Developmental Science*, *20*(3), e12381.

Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., Simson, J., Knox, D., Glowacki, L., & Galbarczyk, A. (2021). Acoustic regularities in infant-directed speech and song across cultures. *BioRxiv*, 2020-04.

Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., Simson, J., Knox, D., Glowacki, L., Galbarczyk, A., Jasienska, G., Ross, C. T., Neff, M. B., Martin, A., Cirelli, L. K., Trehub, S. E., Song, J., Kim, M., Schachner, A., … Mehr, S. A. (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour*, *6*, 1545–1556. https://doi.org/10.1038/s41562-022-01410-x

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1570–1582. https://doi.org/10.1037/0096-1523.26.5.1570

Højen, A., & Nazzi, T. (2016). Vowel bias in Danish word-learning: Processing biases are language-specific. *Developmental Science*, *19*(1), 41–49.

Hsu, H.-C., Fogel, A., & Messinger, D. S. (2001). Infant non-distress vocalization during mother-infant face-to-face interaction: Factors associated with quantitative and qualitative differences☆. *Infant Behavior and Development*, *24*(1), 107–128.

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, *19*(3), 338–352.

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

Kalashnikova, M., Carignan, C., & Burnham, D. (2017). The origins of babytalk: Smiling, teaching or social convergence? *Royal Society Open Science*, *4*(170), 306. https://doi.org/10.1098/rsos.170306

Kaplan, P. S., Bachorowski, J., Smoski, M. J., & Zinser, M. (2001). Role of clinical diagnosis and medication use in effects of maternal depression on infant-directed speech. *Infancy*, *2*(4), 537–548. https://doi.org/10.1207/S15327078IN0204_08

Kidd, E., & Garcia, R. (2022). How diverse is child language acquisition? *First Language*, *42*, 703–735. https://doi.org/10.1177/01427237211066405

Kim, H., Hasegawa-Johnson, M., & Perlman, A. (2011). Vowel contrast and speech intelligibility in dysarthria. *Folia Phoniatrica et Logopaedica*, *63*(4), 187–194.

Kim, H. I., & Johnson, S. P. (2014). Detecting 'infant-directedness' in face and voice. *Developmental Science*, *17*(4), 621–627.

Kitamura, C., & Burnham, D. (1998). *Acoustic and affective qualities of IDS in English*. Fifth international conference on spoken language processing. http://www.isca-speech.org/archive

Kitamura, C., & Lam, C. (2009). Age-specific preferences for infant-directed affective intent. *Infancy*, *14*(1), 77–100. https://doi.org/10.1080/15250000802569777

Kjærbæk, L., & Basbøll, H. (2016). Interaction between input frequency, transparency and productivity in acquisition of noun plural inflection in Danish. *Poznan Studies in Contemporary Linguistics*, *52*(4), 663–686.

Ko, E.-S., Seidl, A., Cristia, A., Reimchen, M., & Soderstrom, M. (2016). Entrainment of prosody in the interaction of mothers with their young children. *Journal of Child Language*, *43*(2), 284–309. https://doi.org/10.1017/S0305000915000410

Kondaurova, M. V., Bergeson, T. R., & Xu, H. (2013). Age-related changes in prosodic features of maternal speech to prelingually deaf infants with cochlear implants. *Infancy*, *18*(5), 825–848. https://doi.org/10.1111/infa.12010

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22), 11850–11857.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*(5326), 684–686. https://doi.org/10.1126/science.277.5326.684

Kvarven, A., Strømland, E., & Johannesson, M. (2020). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behaviour*, *4*(4), 423–434.

Lam, C., & Kitamura, C. (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Developmental Science*, *15*(2), 212–221. https://doi.org/10.1111/j.1467-7687.2011.01118.x

Lam, J., & Tjaden, K. (2013). Acoustic-perceptual relationships in variants of clear speech. *Folia Phoniatrica et Logopaedica*, *65*(3), 148–153.

Lam, J., Tjaden, K., & Wilding, G. (2012). Acoustics of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research*, *55*(6), 1807–1821.

Lam-Cassettari, C., & Kohlhoff, J. (2020). Effect of maternal depression on infant-directed speech to prelinguistic infants: Implications for language development. *PLoS One*, *15*(7), e0236787.

Lavechin, M., Bousbib, R., Bredin, H., Dupoux, E., & Cristia, A. (2020). An open-source voice type classifier for child-centered daylong recordings. *ArXiv Preprint. ArXiv:2005.12656*.

Lee, C. S., Kitamura, C., Burnham, D., & McAngus Todd, N. P. (2014). On the rhythm of infant-versus adult-directed speech in Australian English. *The Journal of the Acoustical Society of America*, *136*(1), 357–365. https://doi.org/10.1121/1.4883479

Lemoine, N. P. (2019). Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, *128*(7), 912–928.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(5), 1431–1436.

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

Liu, H.-M., Tsao, F.-M., & Kuhl, P. K. (2005). The effect of reduced vowel working space on speech intelligibility in mandarin-speaking young adults with cerebral palsy. *The Journal of the Acoustical Society of America*, *117*(6), 3879–3889.

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, *49*(2B), 606–608.

Lovcevic, I., Kalashnikova, M., & Burnham, D. (2020). Acoustic features of infant-directed speech to infants with hearing loss. *The Journal of the Acoustical Society of America*, *148*(6), 3399–3416.

Ludusan, B., Mazuka, R., & Dupoux, E. (2021). Does infant-directed speech help phonetic learning? A machine learning investigation. *Cognitive Science*, *45*(5), e12946.

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), 185–201.

MacWhinney, B. (2014). *The CHILDES project: Tools for analyzing talk, volume II: The database*. Psychology Press. https://doi.org/10.4324/9781315805641
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.
Marklund, E., Marklund, U., & Gustavsson, L. (2021). An association between phonetic complexity of infant vocalizations and parent vowel Hyperarticulation. *Frontiers in Psychology*, *12*, 693866.
Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers speak less clearly to infants than to adults: A comprehensive test of the hyperarticulation hypothesis. *Psychological Science*, *26*(3), 341–347. https://doi.org/10.1177/0956797614562453
Masapollo, M., Polka, L., & Ménard, L. (2016). When infants talk, infants listen: Pre-babbling infants prefer listening to speech with infant vocal properties. *Developmental Science*, *19*(2), 318–328.
McCloy, D. (2016). *phonR: Tools for phoneticians and phonologists*. R package Version 1.0-7. Online: https://cran.r-project.org/web/packages/phonR/phonR.pdf.
McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.
McMurray, B., Aslin, R. N., & Toscano, J. C. (2009). Statistical learning of phonetic categories: Insights from a computational approach. *Developmental Science*, *12*(3), 369–378. https://doi.org/10.1111/j.1467-7687.2009.00822.x
McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition*, *129*(2), 362–378. https://doi.org/10.1016/j.cognition.2013.07.015
Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H., & Mazuka, R. (2017). Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition*, *166*, 84–93. https://doi.org/10.1016/j.cognition.2017.05.003
Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustical Society of America*, *139*(3), 1272–1281. https://doi.org/10.1121/1.4944634
Nearey, T. (1978). Vowel space normalization in synthetic stimuli. *The Journal of the Acoustical Society of America*, *63*(Suppl. 1), S5.
Nguyen, V., Versyp, O., Cox, C. M. M., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the development of turn taking in adult-child vocal interactions. *Child Development.*, *93*, 1181–1200. https://doi.org/10.1111/cdev.13754
Niwano, K., & Sugai, K. (2002). Intonation contour of Japanese maternal infant-directed speech and infant vocal response. *The Japanese Journal of Special Education*, *39*(6), 59–68.
Ohala, J. J. (1980). The acoustic origin of the smile. *The Journal of the Acoustical Society of America*, *68*(Suppl. 1), S33.
Panneton, R., Kitamura, C., Mattock, K., & Burnham, D. (2006). Slow speech enhances younger but not older infants' perception of vocal emotion. *Research in Human Development*, *3*(1), 7–19. https://doi.org/10.1207/s15427617rhd0301_2
Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior and Development*, *14*(4), 415–440. https://doi.org/10.1016/0163-6383(91)90031-M
Parola, A., Simonsen, A., Lin, J. M., Zhou, Y., Huiling, W., Ubukata, S., Koelkebeck, K., Bliksted, V., & Fusaroli, R. (2022). Voice patterns as markers of schizophrenia: Building a cumulative generalizable approach via cross-linguistic and meta-analysis based investigation. *MedRxiv*.
Paulus, M. (2022). Should infant psychology rely on the violation-of-expectation method? Not anymore. *Infant and Child Development*, *31*(1), e2306.
Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science*, *21*(12), 1894–1902. https://doi.org/10.1177/0956797610389189
Peter, V., Kalashnikova, M., Santos, A., & Burnham, D. (2016). Mature neural responses to infant-directed speech but not adult-directed speech in pre-verbal infants. *Scientific Reports*, *6*(1), 1–14. https://doi.org/10.1038/srep34273
Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and speech*, *46*(2-3), 115–154.
Polka, L., Masapollo, M., & Ménard, L. (2022). Setting the stage for speech production: Infants prefer listening to speech sounds with infant vocal resonances. *Journal of Speech, Language, and Hearing Research*, *65*(1), 109–120.
Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychological Science*, *26*(4), 490–498.
Porritt, L. L., Zinser, M. C., Bachorowski, J.-A., & Kaplan, P. S. (2014). Depression diagnoses and fundamental frequency-based acoustic cues in maternal infant-directed speech. *Language Learning and Development*, *10*(1), 51–67.
R Core Team. (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. https://www.R-project.org/
Raneri, D. P. (2015). *Infant-directed speech: Maternal pitch variability, rate of speech, and child language outcomes*. ProQuest Dissertations Publishing.
Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., & Casillas, M. (2021). ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behavior Research Methods*, *53*(2), 818–835. https://doi.org/10.3758/s13428-020-01460-x
Rast, P., & Ferrer, E. (2018). A mixed-effects location scale model for dyadic interactions. *Multivariate Behavioral Research*, *53*(5), 756–775.
Rattanasone, N. X., Burnham, D., & Reilly, R. G. (2013). Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age. *Journal of Phonetics*, *41*(5), 332–343. https://doi.org/10.1016/j.wocn.2013.06.001
Ravanelli, M., & Bengio, Y. (2018). *Speaker recognition from raw waveform with sincnet*. IEEE.
Raviv, L., Lupyan, G., & Green, S. C. (2022). How variability shapes learning and generalization. *Trends in Cognitive Sciences*, *26*(6), 462–483.
Rosslund, A., Mayor, J., Óturai, G., & Kartushina, N. (2022). Parents' hyper-pitch and low vowel category variability in infant-directed speech are associated with 18-month-old toddlers' expressive vocabulary. *Language Development Research*, *2*(1), 223–267.
Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*(2), 339–349. https://doi.org/10.1111/j.1467-7687.2008.00786.x
Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, *15*(6), 608–635. https://doi.org/10.1111/j.1532-7078.2010.00033.x
RStudio Team. (2020). *RStudio: Integrated development for R*. RStudio, PBC. http://www.rstudio.com/
Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*, 181–203. https://doi.org/10.1146/annurev-psych-122216-011805
Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., Laznik, M.-C., & Cohen, D. (2013). Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *PLoS One*, *8*(10), e78103.
Searl, J., & Evitts, P. M. (2013). Tongue–palate contact pressure, oral air pressure, and acoustics of clear speech. *Journal of Speech, Language, and Hearing Research*, *56*(3), 826–839.
Segal, J., & Newman, R. S. (2015). Infant preferences for structural and prosodic properties of infant-directed speech in the second year of life. *Infancy*, *20*(3), 339–351.

Seidl, A., & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, *11*(4), 596–606. https://doi.org/10.1111/j.1467-7687.2008.00704.x

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

Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, *13*(4), 410–420. https://doi.org/10.1080/15250000802188719

Soderstrom, M., Seidl, A., Nelson, D. G. K., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*(2), 249–267. https://doi.org/10.1016/S0749-596X(03)00024-X

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

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

Steen, V. B., & Englund, N. (2021). Child-directed speech in a Norwegian kindergarten setting. *Scandinavian Journal of Educational Research*, *66*, 505–518. https://doi.org/10.1080/00313831.2021.1897873

Steinlen, A. K. (2005). *The influence of consonants on native and non-native vowel production: A cross-linguistic study* (Vol. *30*). Gunter Narr Verlag.

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

Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics*, *27*(1), 24–27.

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

Trecca, F., Bleses, D., Højen, A., Madsen, T. O., & Christiansen, M. H. (2020). When too many vowels impede language processing: An eye-tracking study of Danish-learning children. *Language and Speech*, *63*(4), 898–918.

Trecca, F., Bleses, D., Madsen, T. O., & Christiansen, M. H. (2018). Does sound structure affect word learning? An eye-tracking study of Danish learning toddlers. *Journal of Experimental Child Psychology*, *167*, 180–203.

Trecca, F., McCauley, S. M., Andersen, S. R., Bleses, D., Basbøll, H., Højen, A., Madsen, T. O., Ribu, I. S. B., & Christiansen, M. H. (2019). Segmentation of highly vocalic speech via statistical learning: Initial results from Danish, Norwegian, and English. *Language Learning*, *69*(1), 143–176.

Trecca, F., Tylén, K., Højen, A., & Christiansen, M. H. (2021). Danish as a window onto language processing and learning. *Language Learning*, *71*, 799–833. https://doi.org/10.1111/lang.12450

Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(33), 13273–13278. https://doi.org/10.1073/pnas.0705369104

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vosoughi, S., & Roy, D. K. (2012). A longitudinal study of prosodic exaggeration in child-directed speech. *Proceedings of the Speech Prosody, 6th International Conference*, SProSIG.

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324. https://doi.org/10.1177/0956797614531023

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

Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, *50*(1), 509–535. https://doi.org/10.1146/annurev.psych.50.1.509

Whitfield, J. A., & Goberman, A. M. (2014). Articulatory–acoustic vowel space: Application to clear speech in individuals with Parkinson's disease. *Journal of Communication Disorders*, *51*, 19–28.

Whitfield, J. A., & Goberman, A. M. (2017). Articulatory-acoustic vowel space: Associations between acoustic and perceptual measures of clear speech. *International Journal of Speech-Language Pathology*, *19*(2), 184–194.

Whitfield, J. A., & Mehta, D. D. (2019). Examination of clear speech in Parkinson disease using measures of working vowel space. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2082–2098.

Wieland, E. A., Burnham, E. B., Kondaurova, M., Bergeson, T. R., & Dilley, L. C. (20152015). Vowel space characteristics of speech directed to children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*, *58*(2), 254–267.

Williams, D. R., Zimprich, D. R., & Rast, P. (2019). A Bayesian nonlinear mixed-effects location scale model for learning. *Behavior Research Methods*, *51*(5), 1968–1986.

Woolard, A., Lane, A. E., Campbell, L. E., Whalen, O. M., Swaab, L., Karayanidis, F., Barker, D., Murphy, V., & Benders, T. (2022). Infant and child-directed speech used with infants and children at risk or diagnosed with autism Spectrum disorder: A scoping review. *Review Journal of Autism and Developmental Disorders*, *9*, 290–306. https://doi.org/10.1007/s40489-021-00253-y

Xu, Y., & Chuenwattanapranithi, S. (2007). Perceiving anger and joy in speech through the size code. *Proceedings of the International Conference on Phonetic Sciences*, *16*, 2105–2108.

Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. (2018). Using stacking to average Bayesian predictive distributions (with discussion). *Bayesian Analysis*, *13*(3), 917–1007. https://doi.org/10.1214/17-BA1091

Yellowlees, A., Bursa, F., Fleetwood, K. J., Charlton, S., Hirst, K. J., Sun, R., & Fusco, P. C. (2016). The appropriateness of robust regression in addressing outliers in an anthrax vaccine potency test. *Bioscience*, *66*(1), 63–72.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Cox, C., Dideriksen, C., Keren-Portnoy, T., Roepstorff, A., Christiansen, M. H., & Fusaroli, R. (2023). Infant-directed speech does not always involve exaggerated vowel distinctions: Evidence from Danish. *Child Development*, *00*, 1–25. https://doi.org/10.1111/cdev.13950