Parents' hyper-pitch and low vowel category variability in infant-directed speech are associated with 18-month-old toddlers' expressive vocabulary

Audun Rosslund
University of Oslo, Norway

Julien Mayor
University of Oslo, Norway

Gabriella Óturai
UiT The Arctic University of Norway, Norway

Natalia Kartushina
University of Oslo, Norway

Abstract: The present study examines the acoustic properties of infant-directed speech (IDS) as compared to adult-directed speech (ADS) in Norwegian parents of 18-month-old toddlers, and whether these properties relate to toddlers' expressive vocabulary size. Twenty-one parent-toddler dyads from Tromsø, Northern Norway participated in the study. Parents (16 mothers, 5 fathers), speaking a Northern Norwegian dialect, were recorded in the lab reading a storybook to their toddler (IDS register), and to an experimenter (ADS register). The storybook was designed for the purpose of the study, ensuring identical linguistic contexts across speakers and registers, and multiple representations of each of the nine Norwegian long vowels. We examined both traditionally reported measures of IDS: pitch, pitch range, vowel duration and vowel space expansion, but also novel measures: vowel category variability and vowel category distinctiveness. Our results showed that Norwegian IDS, as compared to ADS, had similar characteristics as in previously reported languages: higher pitch, wider pitch range, longer vowel duration, and expanded vowel space area; in addition, it had more variable vowel categories. Further, parents' hyper-pitch, that is, the within-parent increase in pitch in IDS as compared to ADS, and lower vowel category variability in IDS itself, were related to toddlers' vocabulary. Our results point towards potentially facilitating roles of increase in parents' pitch when talking to their toddlers and of consistency in vowel production in early word learning.

Keywords: infant-directed speech; vocabulary; vowels; language acquisition; Norwegian

Corresponding author(s): Audun Rosslund, Center for Multilingualism in Society across the Lifespan, University of Oslo, Forskningsveien 3A, 0373, Oslo, Norway. Email: audun.rosslund@iln.uio.no

ORCID ID(s): https://orcid.org/0000-0002-2646-8053; https://orcid.org/0000-0001-9827-5421; https://orcid.org/0000-0002-6526-3392; https://orcid.org/0000-0003-4650-5832

Citation: 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. https://doi.org/10.34842/2022.0547

Introduction

When talking to infants and young children, adults fine-tune their speech by slowing it down, heightening their pitch, increasing their pitch range and extending their corner vowels (Fernald, 1989; Kuhl et al., 1997). This speech register, known as infant-directed speech (IDS), functions as a 'perceptual hook' and is suggested to aid infants in the task of language acquisition (Cristia, 2013; Golinkoff et al., 2015). Infants prefer listening to IDS over adult-directed speech (ADS) already two days after birth (Cooper & Aslin, 1990), and this preference increases with language exposure, that is, having stronger effects in older infants, and in infants' native over non-native language (The ManyBabies Consortium, 2020), a preference also correlating with relative language exposure in bilingual infants (Byers-Heinlein et al., 2021). However, there are some inconsistencies in the IDS research, in particular with respect to (1) which properties of IDS may facilitate early language development, (2) whether IDS speech is clearer as compared to ADS, (3) the generalisability of the results to different socio-linguistic contexts, and (4) the methods used to record and analyse IDS. Next, we detail each of these points and describe how they are addressed in the current study.

Both experimental and descriptive studies have reported evidence suggesting that IDS may facilitate language development. Experimental studies have shown that stimuli (words and sentences) that imitate prototypical IDS characteristics facilitate word segmentation (Thiessen et al., 2005), word comprehension (Song et al., 2010) and immediate word learning (Graf Estes & Hurley, 2013; Ma et al., 2011). Analogously, descriptive studies linking properties of parents' IDS to children's language outcomes have found positive correlations between vowel space expansion (larger triangular area between the three corner vowels /i/, /ɑ/, /u/ in IDS as compared to ADS) and expressive vocabulary size (Hartman et al., 2017; Kalashnikova & Burnham, 2018), consonant discrimination (García-Sierra et al., 2021; Kalashnikova & Carreiras, 2021; Liu et al., 2003) and complexity of child vocalizations (Marklund et al., 2021), and between pitch range and expressive vocabulary size (Porritt et al., 2014). Larger vowel space expansion has been hypothesised to increase the clarity of speech, thus making sound categories and words (e.g., *bed vs. bad*) easier to distinguish for language learners. This relationship has originally been observed in adult research on speech perception, when vowel space expansion, together with other phonetic features, were found to lead to better speech intelligibility (see e.g., Garnier et al., 2018), hence clear perceived articulation of speech sounds. Yet, increased vowel space expansion *per* se does not necessarily lead to more intelligible speech (for IDS, see Cristia & Seidl, 2014; Miyazawa et al., 2017). Acoustic analyses of parental recordings revealed increased within-category variability in IDS, which might reduce speech clarity (Cristia & Seidl, 2014; Martin et al., 2015; McMurray et al., 2013; Miyazawa et al., 2017). For example, Japanese mothers of 18–20-month-old toddlers extended their first and second formants when talking to their child, as compared to ADS; yet the increased vowel space area did not lead to more distinct categories due to increased variability in vowel tokens (Miyazawa et al., 2017). Thus, it remains unclear whether the relationship between vowel space expansion in IDS and infants' language outcomes (e.g., Kalashnikova & Burnham, 2018) is attributed to (intentionally) clearer speech provided to the child by the parent, or to a different mediating factor or their combination, such as higher pitch and increased pitch variability, smiling and affect (Benders, 2013), or attempts to appear smaller and less intimidating to the child (Kalashnikova et al., 2017), all of which might potentially lead to vowel space expansion.[1]

Another central question is whether the acoustic properties of IDS – and the potential boosting effect of certain IDS properties in language acquisition – are similar across different socio-linguistic contexts, that is, cultures with varying parenting behaviours, and languages and dialects with varying linguistic structures. As detailed below, this is likely not the case (Saint-Georges et al., 2013, and see e.g., Casillas et al., 2020; Cristia et al., 2022 for descriptions of cultures with infrequent child-directed vocalisations). The majority of studies on IDS have been conducted with American English parents (for the overall prevalence of English in child language studies, see Kidd & Garcia, 2022), who have been described as having more extreme IDS properties than parents in other languages might display (Fernald et al., 1989), questioning the generalizability of the results. While higher and more variable pitch might be the two most robust characteristics of IDS present across most cultures and languages (Broesch & Bryant, 2015; Farran et al., 2016; McClay et al., 2021; Narayan & McDermott, 2016; but see also Han et al., 2020, 2021), vowel space expansion, on the other hand, has not been reported consistently across languages. For instance, increased vowel space expansion in IDS *vs.* ADS has not been found in Dutch (Benders, 2013), German (Audibert & Falk, 2018), Cantonese (Xu Rattanasone et al., 2013), Lenakel and Southwest Tanna (McClay et al., 2021), and reported inconsistently for Norwegian (Englund & Behne, 2006; Kartushina et al., 2021). Further, experimental studies have found that neither British (Floccia et al., 2016) nor German (Schreiner & Mani, 2017) infants segment speech stimuli recorded in natural IDS register in their respective languages, unless these were prosodically exaggerated over and beyond what would be considered 'natural' British and German IDS. Overall, these findings paint the picture that IDS and its potential effect on language development are not uniform, and call for studies of IDS across a wider range of languages and dialectal variations.

[1] We deliberately avoid the term 'hyperarticulation' throughout this manuscript. Although vowel space expansion is, originally, the acoustic proxy for 'hyperarticulation', it is, yet, a component of clear speech; in infant development research, the term is often used interchangeably with clear speech *per se*, not acknowledging potential underlying variability in sound production that may make speech less clear (cf references in the text).

A final concern is the varying procedures used to elicit IDS and to measure its acoustic properties. For example, IDS (and ADS) have been recorded in both home (Narayan & McDermott, 2016) and lab-environments (Benders, 2013), during unstructured (Englund & Behne, 2006) or semi-structured interactions (Kalashnikova & Burnham, 2018), elicited through a picture-description task (Weirich & Simpson, 2019) or a storybook reading (Burnham et al., 2015; McMurray et al., 2013). These differences in the recording contexts can influence the acoustic properties of speech (e.g., Burnham et al., 2015; Miyazawa et al., 2017; Tamis-LeMonda et al., 2017); thus, researchers should weigh the pros and cons of each procedure. In addition, researchers can examine the acoustic properties of parental speech when addressed to their child, the IDS *per se* (e.g., Hartman et al., 2017; Liu et al., 2003; Porritt et al., 2014) or the within-parent difference between the acoustic measures of IDS as compared to ADS, meaning that parents function as their own baseline (e.g., Kalashnikova & Burnham, 2018; Kalashnikova & Carreiras, 2021). Given that these two lines of research in fact capture two complementary constructs of parents' speech – the acoustic features of IDS, and the acoustic difference between the two registers (or the perceived 'adaptation', whether parents modulate it, consciously or not) – there is a need for integrative studies that combine both approaches and examine their respective contribution to the child's early language development.

Hence, the aims of this study were three-fold. First, we sought to assess IDS in comparison to ADS in Norwegian parents speaking a Northern Norwegian dialect. To elicit IDS, we designed a child-friendly storybook[2] (see Methods for details) that enabled us to collect 10 vowel tokens, varying in surrounding consonantal context (5 types), for each of the 9 Norwegian long vowels, providing a more comprehensive analysis of vowels addressed to the child, as compared to describing the three 'corner' vowels in previous research (as also criticised by e.g., Englund, 2018). Parents read this book to their 18-month-old toddler (IDS), as well as to another adult (ADS). This procedure ensured that elicited speech was sampled from identical linguistic contexts across the two registers and speakers (Steinlen & Bohn, 1999; Wang et al., 2015), providing better generalizability across the registers. We examined the acoustic measures of speech that are traditionally reported: that is, pitch, pitch range, vowel duration and vowel space area (Fernald, 1989; Kuhl et al., 1997; Wang et al., 2015), but also novel measures of vowel category variability and vowel category distinctiveness, providing novel proxies/indices for the clarity of speech, as an increased vowel space might also

[2] Note that, for the sake of simplicity, we refer to the storybook-elicited speech read to a child and to an adult as IDS and ADS, respectively.

contain more variability within each vowel category and, hence, lead to less distinct vowel categories (see e.g., Cristia & Seidl, 2014; McMurray et al., 2013). Second, we aimed to evaluate whether the within-parent differences – or adaptation – between IDS and ADS, if any, predicted the expressive vocabulary size of their 18-month-old toddlers (similarly to e.g., Kalashnikova & Burnham, 2018). Finally, we sought to assess whether any of the acoustic measures examined in the current study for IDS, *not* the difference between registers, or adaptation, predicted toddlers' expressive vocabulary (similarly to e.g., Hartman et al., 2017). It is noteworthy that Norwegian language uses vowel formants, vowel length and pitch accent as cues to mark lexical meaning. In addition, Norway is characterised by its dialect diversity, with differences in lexicons, phonemic realisation, and pitch accent patterns across dialects (Mæhlum & Røyneland, 2012). Given that the current knowledge about IDS in Norwegian comes from speakers of the Central Norwegian dialect (Englund, 2018; Englund & Behne, 2005, 2006), the current study (with speakers of the Northern Norwegian dialect) may also highlight potential diversity of IDS in a more fine-grained manner, that is, within-language, but across-dialect.[3]

For our first aim, and in line with previous studies, we expected, as per pre-registration (https://osf.io/7st6w/), that when addressing speech to their child (IDS), in comparison to an adult (ADS), Norwegian parents will produce: higher pitch, wider pitch range and increased vowel duration. With respect to the vowel space area, Englund & Behne (2006) found a decrease in Norwegian parents' IDS addressed to 1–6-month-old infants, whereas Kartushina and colleagues (2021) found an increase in Norwegian parents' IDS addressed to 8-month-old infants. These differences in vowel space can be due to either children's ages (0–6-month-olds *vs.* 8-month-olds), differences in dialects (Central *vs.* Eastern Norwegian), or methods to compute vowel space (using /ɑ:/ *vs.* /æ:/ as the extreme/corner open vowel in Englund & Behne, 2006 and Kartushina et al., 2021, respectively), or a combination of these factors. Given that the current study examined parents speaking a Northern Norwegian dialect directed to older toddlers, and measured vowel space using the /æ:/ vowel as the most extreme open vowel in Norwegian, we predicted, in line with Kartushina and colleagues (2021), vowel space expansion in IDS, as compared to ADS. Finally, in line with recent results in Norwegian (Kartushina et al., 2021), English (Cristia & Seidl, 2014; McMurray et al.,

[3] We note that distinguishing dialects from languages is not necessarily linguistically meaningful, as this distinction is primarily linked to political and cultural factors (yet, for a recent attempt, see Wichmann, 2020).

2013) and Japanese (Martin et al., 2015; Miyazawa et al., 2014), we expected vowel categories to be less compact and less distinct in IDS, as compared to ADS.

For our second aim, to evaluate whether the within-parent differences – or adaptation – between IDS and ADS, if any, predict the vocabulary of their toddlers, in line with previous research, we expected that increases in pitch, pitch range and vowel duration would be positively related to toddlers' expressive vocabulary. Given that pitch accent and vowel duration are lexically meaningful cues in Norwegian (they are used to distinguish words, as, for example in *tak* [roof] *vs. takk* [thanks] or *bønder* [farmer] *vs. bønner* [beans]), we expected that toddlers would benefit from input that emphasises these cues in IDS, especially since, at 18 months of age, their expressive vocabulary is rapidly increasing. In addition, we expected a positive relationship between vowel space expansion and toddlers' expressive vocabulary (as found in Hartman et al., 2017; Kalashnikova & Burnham, 2018). Finally, as we expected increased within-vowel category variability and less between-vowel distinctiveness in IDS, as compared to ADS (e.g., Cristia & Seidl, 2014; McMurray et al., 2013), we anticipated that parents who produce less variable and/or more distinct vowel categories would, by means of facilitating speech sound discrimination and representations, boost their child's word learning. Hence, we expected a negative relationship between vowel category variability and toddlers' expressive vocabulary, but a positive relationship between vowel category distinctiveness and toddlers' expressive vocabulary. To summarise, we hypothesise that the 'ideal' IDS adaptation benefiting early word learning contains exaggerated (a) pitch and pitch range, (b) vowel duration, and (c) vowel space, and (d) precise vowel tokens with (e) little variability within each category.

Last, and for our third aim, we assessed whether any of the acoustic measures examined in the current study for parents' IDS itself, *not* the difference between the registers, predicted toddlers' vocabulary, and we expected that the same acoustic features as those that were emphasised in IDS when compared to ADS (within-parent differences between the registers), would be associated with toddlers' expressive vocabulary. That is, parent-specific pitch, pitch range, vowel duration, vowel space area and vowel category distinctiveness in IDS would be positively related to toddlers' expressive vocabulary, while vowel category variability would be negatively related to toddlers' expressive vocabulary.

Method

Participants

Twenty-one parent-toddler dyads from the city of Tromsø (Northern Norway) participated in the current study. Two additional dyads were recruited, but excluded from the analysis, due to missing audio files (n = 1) and less than 75% exposure to Norwegian (n = 1). For the final sample, all parents (16 mothers, 5 fathers) were native speakers of Norwegian, raised in Northern Norway and spoke the Northern Norwegian dialect. All parents cohabited with their toddlers and the toddlers' other parent, and reported to provide at least 50% of speech input to their toddler as compared to the other parent. Toddlers (9 girls, 12 boys, *M* age = 17.9 months, *SD* = 0.43) were exposed, on average, to 97.5% of Norwegian (*SD* = 7.49) and none had reported any visual or auditive impairments.[4] Socioeconomic status (SES), reported as mother's highest education level, ranged from 1 (secondary school) to 5 (doctoral degree), with the median being 3 (bachelor's degree).

Data collection took place in the BabyLab at the Department of Psychology, University of Tromsø. After receiving invitations through advertisement on social media, at the university, local library or health station, parents who agreed to participate with their child in the study signed an informed consent form, and within the five days after their visit to the lab, answered a web questionnaire that included general demographic questions and questions about their toddlers' linguistic environment. The online questionnaire included the Norwegian adaptation of the MacArthur-Bates Communicative Development Inventories (CDI) –Words and Sentences form (Simonsen et al., 2014). Individual raw CDI scores (the number of words that parents reported their child to produce) were converted to daily percentiles using the normative Norwegian data from Wordbank (Frank et al., 2017; for the conversion procedure, see Kartushina al., 2022); the mean score was 37.6 (*SD* = 29.3, range = 1–93). The current study was conducted according to the guidelines laid in the Declaration of Helsinki, with written informed consent obtained from a parent or a guardian for a child before any assessment or data collection. The study has been approved by the Norwegian Centre for Research Data (NSD, ref. 56312), and the local ethical committee at the Department of Psychology, University of Oslo. The pre-registration, data, stimuli and analysis script for the study are openly available at the Open Science Framework (OSF) project's page (https://osf.io/7st6w/).

[4] Two of the toddlers were reported to be born 'too early'. The exclusion criteria for toddlers was to be born before 37 weeks of gestation (i.e., premature according to medical convention). However, poor wording of this specific question in our questionnaire made parents' responses ambiguous. The wording of the question was open, not specific to the number of weeks and did not include the term 'premature'. Thus, we were not able to know whether these two toddlers were in fact premature or simply born any time (e.g., one or two days) before the expected due date. Comparing these two toddlers to the rest of the sample on the key measures did not reveal any differences (see Appendix 2). We, therefore, included them for the analyses.

Procedure and Stimuli

Upon the arrival to the BabyLab, parents and their toddlers were familiarised with the lab environment and experimenters and received information about the course of their visit. Seven of the toddlers took part in an unrelated experiment on motor imitation prior to the recordings. Parents were not aware of the specific purpose of the study, or which parts of their recorded speech were of interest to the researchers, until after they had completed the recording sessions.

The IDS and ADS recordings took place either in the waiting area of the BabyLab, or in an adjacent child-friendly room. Both IDS and ADS were elicited from the parent through reading a child-friendly storybook, specifically created for the purpose of the study. The storybook was written in Norwegian Bokmål[5] and consisted of five pages, 39 sentences and 327 words. Each page had a colourful illustration and a short child-friendly narrative (Table 1); the narratives were not connected with each other. The nine long Norwegian vowels (/ɑː/, /eː/, /iː/, /uː/, /ʉː/, /yː/, /æː/, /øː/, and /ɔː/) were represented by five unique words repeated twice throughout the storybook, for a total of 90 target vowels. The words were mono- and bisyllabic lexical and function words, most of them reported to be known by a large proportion of toddlers at this age (Simonsen et al., 2014). Words were counterbalanced in terms of their position within a sentence, so that each target vowel was present in at least one start-, mid- and end-sentence word. The target vowel was in a stressed position within the word, and, for the bisyllabic words, with the two exceptions, the target vowel was always placed in the first syllable. See Appendix 1 for an overview of target vowels within words.

During the IDS recording, the parent read the storybook to their toddler either sitting on their lap or next to them. Parents were instructed to read and interact with their child as they would typically do when reading a book at home. Parents did not receive any instructions with respect to the dialect to use (recall the book was written in Norwegian Bokmål, which is close to the Eastern, Oslo-area, dialect); all parents chose to read in their Northern Norwegian dialect, that is, adapting the grammatical gender, the phonemic realisation, and the intonation patterns to this dialect. During the ADS recording, parents read the same storybook to the experimenter (a native speaker of Norwegian), with no further instructions but to read the book naturally as if reading to an adult. Again, parents chose to read in their Northern Norwegian dialect. During

[5] Dialects are not used in written text; hence this is one of two official, dialect-neutral, written forms of Norwegian.

the ADS recording, a second experimenter cared for the toddler outside of the parents' field of vision. Due to limited resources, the second experimenter was not available for three parent-toddler dyads. The order of the recordings was counterbalanced; half of the parents started with the IDS, and the other half started with the ADS. All sessions were recorded with an Olympus DS-3000 handheld voice recorder in 16-bit/44.1 kHz. After the recordings, toddlers received a small toy or a book as a token of appreciation.

Table 1. ***Example of text from one page in the storybook (words with target vowels in bold, IPA transcripts in brackets)***

Original	English translation
Mamma-sjiraffen **skjærer** [ʂæ:rer] en **skive** [ʂi:və] av **brødet** [brø:ə]. Den lille sjiraffen ligger på **magen** [mɑ:gən], med den ene **foten** [fu:tən] i været. Han vil heller ha **kake** [kɑ:kə] og **banan** [bɑnɑ:n]. Mamma-sjiraffen **skjærer** [ʂæ:rer] enda en **skive** [ʂi:və] av **brødet** [brø:ə], og legger fram en **skje** [ʂe:] til grøten. "Vi kan **spise** [spi:se] **kake** [kɑ:kə] og **banan** [bɑnɑ:n] etterpå", sier mamma-sjiraffen. "**Bra**! [brɑ:]", sier den lille sjiraffen.	Mommy-giraffe **cuts** a **slice** of **bread**. The little giraffe is lying on his belly, with one **foot** in the air. He would rather have **cake** and **banana**. Mommy-giraffe **cuts** another **slice** of **bread**, and lays out a **spoon** for the porridge. "We can **eat** **cake** and **banana** later", says Mommy-giraffe. "**Great**!", says the little giraffe.

Data Processing and Acoustic Measures

Three trained native speakers of Norwegian listened to the audio recordings in Praat (Boersma & Weenink, 2020) and marked the target speech segments. First, they segmented parents' speech and marked the onset and the offset of the phrases, necessary for the pitch analyses. A phrase was defined as a portion of continuous speech with intact pitch tracks, without interruptions (e.g., interference from the child), enclosed by approximately 500 ms of silence, typically a pause where the parent drew breath. In other words, the length and the content of a phrase varied across segments and could include short utterances as well as full sentences. In total, we identified 923 phrases in IDS and 818 phrases in ADS. A customised Praat script (Hirst, 2012) automatically extracted the duration and the minimum, maximum, and mean pitch (F0) in Hz for each phrase. 133 phrases (7.6%) were manually corrected due to errors in the octave jumps (i.e., pitch tracks printed one octave higher than intended). As pitch perception follows a logarithmic scale, all Hz values were converted to semitones using the following formula *semitones =12*log^2(F0/constant)*, as in Kalashnikova & Burnham (2018), with 10 as a constant (i.e., semitones-above-10-hertz). Pitch range was computed as the difference between the minimum and the maximum pitch value (in semitones) within each phrase.

Second, we identified and manually annotated the target vowels. Only audible target vowels, with a minimum length of 30 ms, with no noise and with visually trackable first (F1) and second (F2) formants were segmented. We followed the same vowel onset and offset boundary definition as in Cristia & Seidl (2014). In total, we identified 1577 vowels in IDS and 1527 vowels in ADS. A customised Praat script (Hirst, 2012) was run to collect vowel duration (in ms) and the mean F0, F1 and F2 (in Hz), with the pre-specified formant ceiling values at 5500 Hz for mothers and 5000 Hz for fathers. 297 vowel segments (9.6%) were manually corrected due to errors in the formant estimates (typically identifying F1 as F2, or F3 as F2, which could be due to high F0, see Monsen & Engebretson, 1983). See Figure 1 and Table 2 for an overview of all vowel segments. Computations of the different vowel-based measures are explained below.

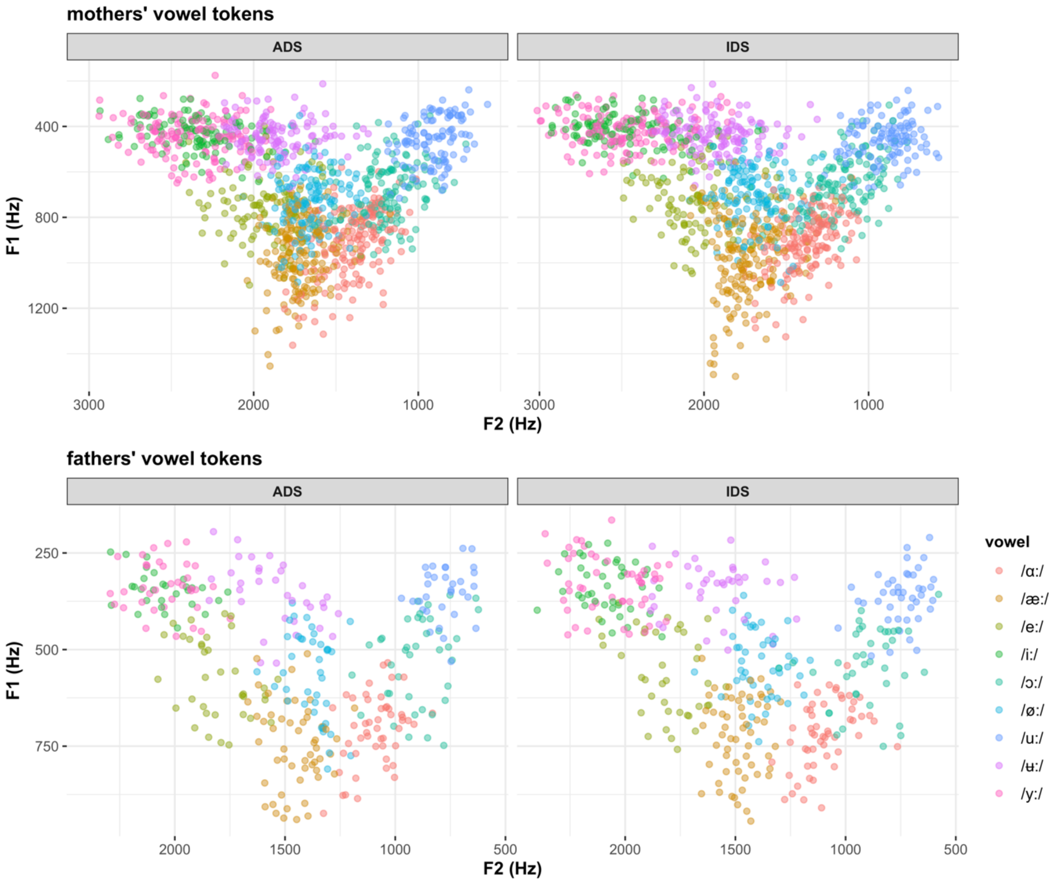


Figure 1. ***Mother's and father's vowel tokens in F1-F2 space by register***

Table 2. ***Number of tokens, mean duration (ms) and mean formant frequencies (Hz) for each target vowel across IDS and ADS registers for mothers and fathers, with standard deviations in parentheses***

	ADS								**IDS**							
	mothers				fathers				mothers				fathers			
	n	duration	F1	F2	n	duration	F1	F2	n	duration	F1	F2	n	duration	F1	F2
/iː/	107	96.0	434	2350	29	123	529	896	118	114	417	2460	40	116	338	2080
		(30.1)	(71.9)	(244)		(43.7)	(112)	(132)		(58.9)	(69.9)	(272)		(48.2)	(63.9)	(129)
/yː/	132	102	442	2390	36	120	340	2040	128	128	415	2470	48	106	331	2070
		(29.9)	(88.7)	(227)		(35.8)	(64.8)	(131)		(39.3)	(75.5)	(265)		(31.0)	(74.9)	(159)
/eː/	97	109	747	1960	36	99.5	576	1800	92	120	734	2080	38	107	564	1820
		(45.3)	(135)	(192)		(35.8)	(101)	(140)		(50.2)	(139)	(191)		(35.2)	(108)	(144)
/øː/	137	111	728	1660	40	117	566	1380	130	121	701	1690	46	113	539	1390
		(36.0)	(130)	(140)		(27.1)	(113)	(89.7)		(40.5)	(127)	(159)		(21.1)	(85.3)	(104)
/æː/	183	106	951	1710	53	113	747	1450	176	119	974	1730	58	115	725	1480
		(37.8)	(147)	(135)		(33.3)	(108)	(109)		(50.3)	(182)	(167)		(44.7)	(102)	(91.3)
/ʉː/	123	108	448	1860	34	102	358	1550	131	127	432	1920	41	103	349	1580
		(44.4)	(80.4)	(196)		(30.7)	(82.8)	(164)		(53.1)	(78.1)	(209)		(25.8)	(70.5)	(181)
/uː/	110	133	461	909	32	119	353	775	113	169	462	894	42	143	364	747
		(55.7)	(105)	(133)		(30.4)	(67.6)	(89.7)		(86.5)	(89.5)	(152)		(60.2)	(68.9)	(102)
/ɔː/	128	112	683	1190	39	122	529	896	126	140	656	1180	39	125	542	902
		(38.0)	(130)	(169)		(43.7)	(112)	(132)		(60.3)	(125)	(201)		(36.7)	(98.7)	(149)
/ɑː/	162	138	911	1400	49	135	707	1090	157	171	907	1380	54	154	717	1090
		(57.8)	(153)	(165)		(38.8)	(90.2)	(104)		(89.9)	(131)	(151)		(67.1)	(86.5)	(113)
Mean	131	113	672	1700	38.7	117	529	1420	130	135	654	1740	45.1	121	510	1460
	(27.3)	(44.6)	(237)	(481)	(7.8)	(36.7)	(181)	(435)	(24.4)	(64.9)	(246)	(533)	(7.01)	(47.2)	(178)	(462)

Vowel Space Area

For the vowel space area (VSA), we measured the overall size of the F1-F2 vowel space (in Hz^2) with the *phonR* package (McCloy, 2016), using the average F1 and F2 (in Hz) for each vowel category and the following formula (exemplified with three vowels, where 'ABS' is the absolute value): $ABS\ ½ \times [(F1/vowel_1/ \times (F2/vowel_2/ - F2/vowel_3/) + F1/vowel_2/ \times (F2/vowel_3/ - F2/vowel_1/) + F1/vowel_3/ \times (F2/vowel_1/ - F2/vowel_2/)]$ and so forth, previously used in IDS research (Kalashnikova & Burnham, 2018; Kuhl et al., 1997; Liu et al., 2003). For each register and each parent, we computed three different vowel space area (VSA) measures: one using the corner vowels /iː/, /**ɑ**ː/, /uː/ ("VSA_a"), in line with previous research in IDS, including Norwegian (Englund, 2018); one using the corner vowels /iː/, /**æ**ː/, /uː/ ("VSA_æ"), as, based on earlier findings in Norwegian (Kartushina et al., 2021) and also confirmed by our data, /æː/ is the most extreme Norwegian open vowel in the F1-F2 space (see Figure 1). In addition, we computed a measure of vowel space area including all border vowels; /ɑː/, /eː/, /iː/, /uː/, /ʉː/, /æː/, /ɔː/ ("VSA_full"), as this would measure most accurately the total vowel area, as the actual vowel space may not necessarily be accurately represented by a triangle.

Vowel Category Variability

The vowel category variability score is an index of the within-category precision in vowel production.[6] The variability of each vowel category in the F1-F2 vowel space (as also used by Hartman and colleagues, 2017) was measured by fitting F1 and F2 (Hz) of all vowel tokens, exemplifying the category, to a customised MatLab script (Kartushina & Frauenfelder, 2014), which calculated the area of an ellipse (Hz^2) for each vowel category, participant, and register, with the following formula: *ellipse_area* = $F1 \times \ F2 \times \pi$, where σF1 is 1 standard deviation of the mean of F1, and σF2 is 1 standard deviation of the mean of F2. Since the distribution of the productions in F1/F2 space was assumed to be elliptical, we estimated the angles of the major and minor axes of an ellipse centered on the mean of the productions (in order to determine the orientation of the axes). Therefore, a low vowel category variability score indicated more compact vowel categories, whereas a high vowel category variability score indicated looser vowel categories.

[6] Note that in the pre-registration, we referred to this measure as 'vowel category compactness'.

Vowel Category Distinctiveness

For vowel category distinctiveness, we measured how distinct participants' vowel categories were from each other in the F1-F2 vowel space. Thus, while vowel category variability indicates the precision of vowel production within each category, vowel category distinctiveness indicates the discriminability of the categories, i.e., the degree of overlap, taking into account their distribution within the full vowel space. Vowel category distinctiveness was computed as the between-vowel category Sum of Squares (the squared distances of category cluster centroids from the overall vowel space centroid) divided by the total Sum of Squares (squared distances of individual vowel tokens from the overall vowel space centroid), for each participant and register, for 8 vowel categories (we omitted the category /y/, as it fully overlaps with the Norwegian /i/ in the F1-F2 space, as the distinguishing feature is F3). See Appendices 3A and 3B for a thorough explanation and visual representation of the measure as a function of the amount of overlap between the vowel categories. Thus, vowel category distinctiveness can be thought of as a clustering performance quotient, indexing the proportion of variance in F1 and F2 explained by the vowel category identity, ranging from 0 (cluster/category membership explains no variance) to 1 (cluster/category membership explains all variance). In sum, with these three F1-F2 based measures, computed across vowel categories, we aimed to thoroughly describe the distinguishing features of parents' vowel production in IDS. For further details on the computation of measures, we refer readers to the available code on the OSF project page (https://osf.io/7st6w/).

Results

The results are structured according to the three aims of the current study; 1) to examine whether there were differences in acoustic properties, both traditional (pitch, pitch range, vowel duration and vowel space area) and novel (vowel category variability and vowel category distinctiveness), between IDS and ADS, 2) to assess the role of within-parent differences between the IDS and ADS registers in predicting toddlers' expressive vocabulary, and 3) to assess the role of acoustic properties of IDS in predicting toddlers' expressive vocabulary. All analyses were preregistered and conducted in R (R Core Team, 2020), with libraries and their versions listed in Appendix 4.

Acoustic Properties of IDS and ADS

Between-register differences in the acoustic measures were assessed with a linear mixed-effect model separately for each acoustic measure. The fixed structure was similar for all models and included register, parent gender and their interaction; the random structure included participant, as well as register and vowel category for some models (cf details below). Models were fitted with the *lme4* package (Bates et al.,

2015) and the model assumptions, including normality and homogeneity of residuals, were visually inspected on diagnostics plots derived from the check_model() function from the *performance* package (Lüdecke et al., 2021). Models were analysed with the Anova() function from the *car* package (Fox & Weisberg, 2018) with the p-values obtained from the *lmerTest* package, using Satterthwaite approximation (Kuznetsova et al., 2017). All model results are shown in Table 3, and between-register differences are visualised in Figure 2.

Pitch

As shown in Table 3, there was a significant effect of register and parent gender on pitch. That is, as expected, parents had a higher mean pitch (all reported in semitones) in IDS (M = 54.1, SD = 6.67) than in ADS (M = 51.4, SD = 6.02), Hedges g = 1.28. Further, mothers had overall higher mean pitch (M = 55.8, SD = 3.51) than fathers (M = 43.9, SD = 4.93). The register by parent gender interaction was not significant.

Pitch Range

As shown in Table 3, there was a significant effect of register on pitch range: As expected, parents had a wider pitch range (all reported in semitones) in IDS (M = 14.6, SD = 6.39) than in ADS (M = 13.3, SD = 5.59), Hedges g = 0.44. The main effect of parent gender on pitch range, and the interaction effect of parent gender and register were not significant.

Vowel Duration

As shown in Table 3, there was a significant effect of register on vowel duration. Note that we log-transformed the outcome measure in the linear mixed-effects model, because the initial model violated the assumption of normality of residuals (see pre/post diagnostics plots in Appendix 5A and 5B). That is, as expected, parents produced longer vowels (reported in ms here for ease of interpretation) in IDS (M = 131, SD = 61.1) than in ADS (M = 114, SD = 43), Hedges g = 1.05. However, as can be seen in the follow-up analyses using *lsmeans* (Lenth, 2016), the main effect is due to the mothers prolonging their vowels to a greater degree in IDS (M = 135, SD = 64.9) as compared to ADS (M = 113, SD = 44.6, $t(16.4) = -5.7$, p = < .001), whereas fathers' vowel duration did not differ significantly between the registers (IDS: M = 121, SD = 47.2, ADS: M = 117, SD = 36.7, $t(19.6) = -0.3$, $p = .766$).

Vowel Space Area

As shown in Table 3, there was a significant effect of register on all three of our vowel space area measures. To facilitate the descriptive statistics, vowel space areas (reported in Hz^2) were divided by 1000, hence, kHz^2. As expected, parents expanded their vowel space area in IDS (VSA_a: *M* = 339, *SD* = 99.7; VSA_æ: *M* = 379, *SD* = 124; VSA_full: *M* = 441, *SD* = 113) as compared to ADS (VSA_a: *M* = 303, *SD* = 104; VSA_æ: *M* = 335, *SD* = 106; VSA_full: *M* = 389, *SD* = 120), Hedges *g* = 0.58; 0.55; 0.54, for VSA_a, VSA_æ and VSA_full, respectively. Further, for all vowel space area measures, mothers had overall larger vowel space areas (VSA_a: *M* = 349, *SD* = 97.7; VSA_æ: *M* = 389, *SD* = 44.4; VSA_full: *M* = 445, *SD* = 112) than fathers (VSA_a: *M* = 232, *SD* = 46.4; VSA_æ: *M* = 253, *SD* = 11.3; VSA_full: *M* = 322, *SD* = 88.4). The register by parent gender interaction was not significant for any measure of vowel space.

Vowel Category Variability

As shown in Table 3, there was a significant effect of register and parent gender on vowel category variability. Note that we log-transformed the outcome measure in the linear mixed-effects model, because the initial model violated the assumption of normality of residuals (see pre/post diagnostics plots in Appendix 6A and 6B). To facilitate the interpretability of the descriptive statistics, we report the non-log transformed vowel category variability in kHz^2. As expected, parents had more variable categories in IDS (*M* = 311, *SD* = 225) than in ADS (*M* = 273, *SD* = 0205), Hedges *g* = 0.44. Further, mothers had overall more variable categories (*M* = 333, *SD* = 228) than fathers (*M* = 161, *SD* = 80.4). The register by parent gender interaction was not significant.

Vowel Category Distinctiveness

As shown in Table 3, parent gender was the only significant effect on vowel category distinctiveness, with mothers having overall less distinct categories (*M* = 0.88, *SD* = 0.04) than fathers (*M* = 0.93, *SD* = 0.02), Hedges *g* = -1.50. Contrary to our expectation, there were no differences between the two registers, and the register by parent gender interaction was not significant.

Table 3. ***Model outputs on acoustic differences between the IDS and ADS registers (n = 21 parent-toddler dyads)***

Model	Parameter	2	*df*	*p*
Pitch ~	Register	40.72	1	<.001***
Register * Gender +	Gender	127.2	1	<.001***
(1 + Register | Participant)[7]	Register * Gender	2.209	1	.137
Pitch_range ~	Register	4.308	1	.038*
Register * Gender +	Gender	1.016	1	.314
(1 + Register | Participant)[7]	Register * Gender	0.121	1	.728
Vowel_duration ~	Register	25.09	1	<.001***
Register * Gender +	Gender	0.159	1	.690
(1 + Register | Participant) + (1 + Register | Vowel)[8]	Register * Gender	8.020	1	.005**
Vowel_space_a ~	Register	7.559	1	.006**
Register * Gender +	Gender	7.541	1	.006**
(1 | Participant)	Register * Gender	0.638	1	.424
Vowel_space_æ ~	Register	7.351	1	.007**
Register * Gender +	Gender	8.077	1	.004**
(1 | Participant)	Register * Gender	2.389	1	.122
Vowel_space_full ~	Register	6.656	1	.010*
Register * Gender +	Gender	6.298	1	.012*
(1 | Participant)	Register * Gender	0.982	1	.322
Vowel_category_variability ~	Register	8.891	1	.003**
Register * Gender +	Gender	7.700	1	.006**
(1 | Participant) + (1 + Register | Vowel)[8]	Register * Gender	0.203	1	.652
Vowel_category_distinctiveness ~	Register	0.001	1	.977
Register * Gender +	Gender	9.683	1	.002**
(1 | Participant)	Register * Gender	0.067	1	.796

*p < .05, **p < .01, ***p < .001

[7] Note that these models deviate from that specified in the pre-registration, where we included a random structure of the segmented phrase in which we extracted the pitch tracks. Given that the number of phrases and their content varied across registers, it was impossible to have similar segment structures.

[8] Recall that the outcome variable was log-transformed.

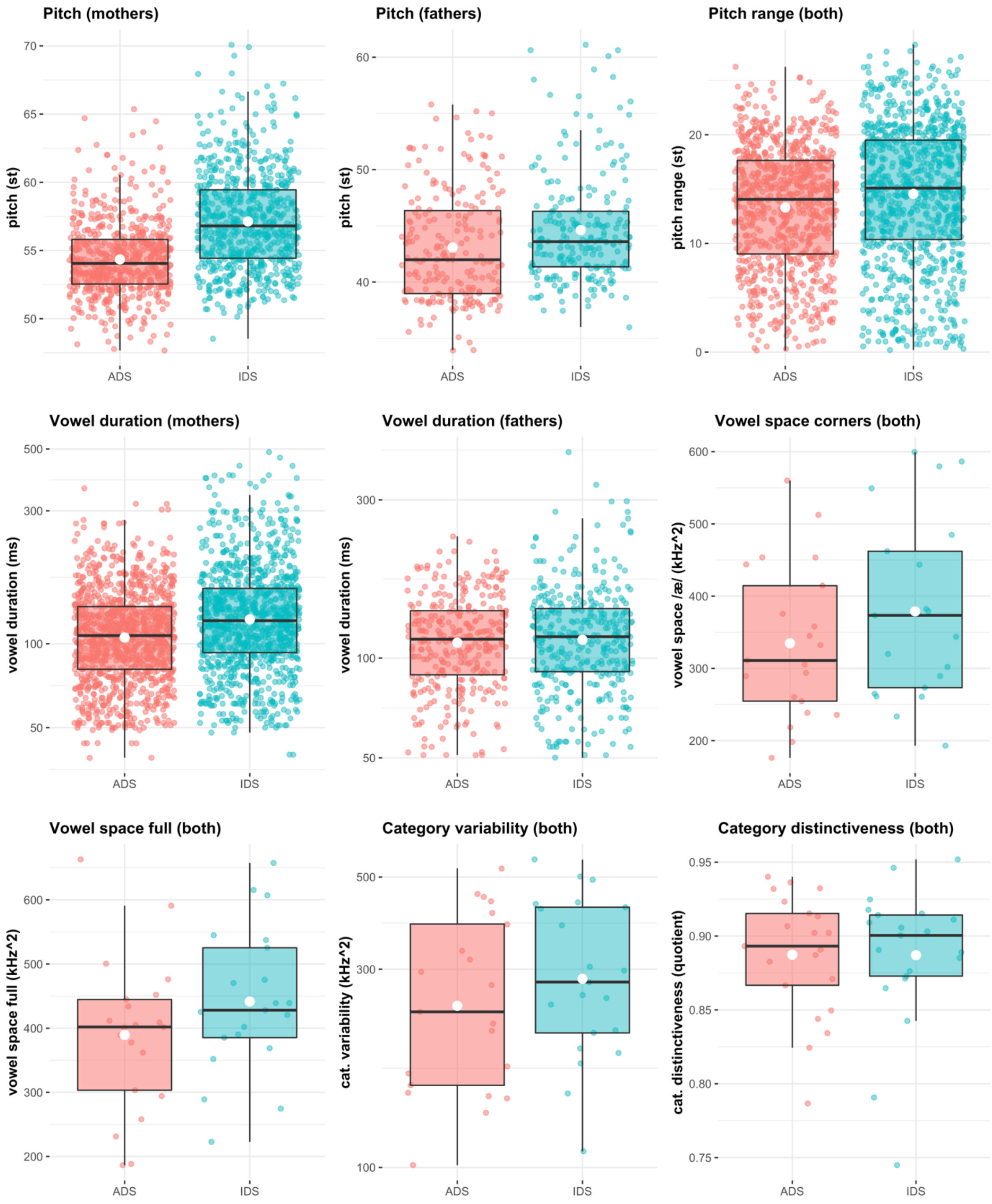


Figure 2. ***Boxplots of acoustic measures in IDS and ADS. Note that the white dots represent the mean. Pitch and vowel duration are visualised separately for parent gender. For vowel duration and category variability, y-axis ticks indicate the scale in the original units, but data is plotted with log-transformed units as this was used in our models. Pitch and pitch range are in semitones, vowel duration in milliseconds, vowel spaces and category variability in kHz² and category distinctiveness in quotients.***

Within-Parent Differences Between IDS and ADS and Toddlers' Expressive Vocabulary

To assess whether the differences parents may have in IDS as compared to ADS predicted toddlers' vocabulary, we computed, first, the ratio between the registers for all the examined acoustic measures, by dividing, for each parent, the average IDS measure by the respective average ADS measure. One exception to this was the vowel space measures – as there was only one measure per register, we did not have to compute the average. A score above 1 indicated a higher value of a specific acoustic measure in IDS, that is, a *hyper*-feature in IDS, and a score below 1 indicated a higher value of a specific acoustic measure in ADS, that is, a *hypo*-feature in IDS. Next, we z-transformed these ratios for each acoustic measure, to facilitate model convergence. Finally, we fitted a beta-regression model using the *betareg* package (Cribari-Neto & Zeileis, 2010), with the outcome measure toddlers' CDI percentiles divided by 100, as required for the beta distributions. The model parameters were:[9]

> CDI percentile ~ Pitch diff_z + Pitch range diff_z + Vowel duration diff_z + Vowel space_æ diff_z + Vowel space_full diff_z + Vowel category variability diff_z

As can be seen in the model output (produced by the *summary* function on the model) reported in Table 4, parents' pitch difference significantly predicted toddlers' vocabulary in percentiles, whereas the other acoustic measures did not. As visualised in Figure 3, parents' hyper-pitch, i.e., an increase in IDS as compared to ADS, was positively related to vocabulary, that is, CDI percentiles increased by 0.71 when pitch difference increased by one standard deviation of the sample mean with all other factors kept at an average. To examine if such an increase in pitch was a deliberate choice parents made, we computed, in an exploratorily analysis, a correlation between parents' hyper-pitch and a mean score of four items retrieved from our background

---

[9] Given that some of our acoustic measures were highly correlated, such as the two measures of vowel space (using corner vowels versus using the full vowel space), and vowel category variability and vowel category distinctiveness (see Appendix 7), we used the variance inflation factor (VIF) to estimate multicollinearity between the predictors. We took a conservative approach and kept predictors within the VIF < 2.5 (e.g., Zuur et al., 2010). Fitting the pre-registered model resulted in high VIFs for the vowel category variability (VIF = 3.01) and vowel category distinctiveness (VIF = 3.29), and so we excluded the latter, given that we did not find any differences between parents' category distinctiveness across registers.

questionnaire that examined parental attitudes towards early language development, developed in Frank and Hembacher (2020)[10], finding no significant relationship, $r_s(19) = .25$, $p = .275$, suggesting that parents' variability in hyper-pitch in IDS was not related to their differences in beliefs that parents need to provide salient linguistic input in an infant-friendly manner to their child.

Table 4. ***Beta-regression results for vocabulary by parents' difference in IDS vs. ADS (n = 21 parent toddler dyads)***

Parameter	*estimate*	*SE*	z	*p*
Intercept	-0.500	0.203	-2.464	.014*
Pitch diff_z	0.705	0.235	3.008	.003**
Pitch range diff_z	0.313	0.237	1.320	.187
Vowel duration diff_z	0.124	0.235	0.529	.597
Vowel space_æ diff_z	0.564	0.317	1.780	.075
Vowel space_full diff_z	-0.325	0.303	-1.075	.283
Vowel category variability diff_z	0.406	0.230	1.766	.077

$^{*}p < .05$, $^{**}p < .01$

Acoustic Properties of Parents' IDS and Toddlers' Expressive Vocabulary

Finally, to assess whether the acoustic properties of parental input in IDS predicted toddlers' vocabulary, independently of any differences between the IDS and ADS registers, we z-transformed mean values on all our acoustic measures in IDS, separately for mothers and fathers. Given that there are physical differences between males and females impacting the acoustics of speech, this was necessary so that, for example, lower pitch and smaller vowel spaces in fathers would not cloud any results. This approach is a deviation from our pre-registered pipeline, where we suggested, 1) to run the model with mothers only, 2) to transform F1 and F2 from Hz to Bark to normalise, then recompute vowel-based measures. The latter did not seem to adjust for between-gender differences as well as predicted. Hence, we chose to instead standardise

[10] The items were the following statements (responses indicating level of agreement on a 0-6 scale): 'Parents can help babies learn language by talking to them' / 'When speaking to a young child, I often speak slower and more clearly' / 'Reading books to children is not useful until they have learned to speak' (reverse coded) / 'When speaking to a young child, I often use a different voice with a more lively tone.'

measures within each gender group. As before, we fitted and analysed a beta-regression model with toddlers' CDI percentiles divided by 100 as our outcome measure. The model parameters were:[11]

> CDI percentile ~ Pitch IDS_z + Pitch range IDS_z + Vowel duration IDS_z + Vowel space_full IDS_z + Vowel category variability IDS_z

The model output can be seen in Table 5. Vowel category variability in IDS significantly predicted toddlers' vocabulary size, whereas the other acoustic measures were not significant. As visualised in Figure 4, parents with more variable vowel categories in IDS had toddlers with lower vocabulary sizes (in percentiles), that is, CDI percentiles decreased by 0.50 when the vowel category variability increased by one standard deviation of the sample mean with the other factors being kept at an average. As a complementary analysis, we provide a correlation matrix and a correlation network plot with all acoustic measures in Appendix 7A and 7B.

Table 5. ***Beta-regression results for vocabulary by parents' input in IDS (n = 21 parent-toddler dyads)***

Parameter	*estimate*	*SE*	z	*p*
Intercept	-0.487	0.225	-2.168	.030*
Pitch_IDS_z	0.266	0.256	1.040	.298
Pitch range_IDS_z	0.051	0.254	0.202	.840
Vowel duration_IDS_z	-0.160	0.274	-0.585	.559
Vowel space_full_IDS_z	0.296	0.264	2.121	.262
Vowel category variability_IDS_z	-0.499	0.254	-1.962	.050*

*$p < .05$

[11] Fitting the pre-registered model resulted in high VIFs for vowel category distinctiveness (VIF = 2.96), vowel space_æ (VIF = 8.34) and vowel space_full (VIF = 9.62). We chose to keep the latter of the vowel space measures, given that this would maximise the information about parents' vowel space.

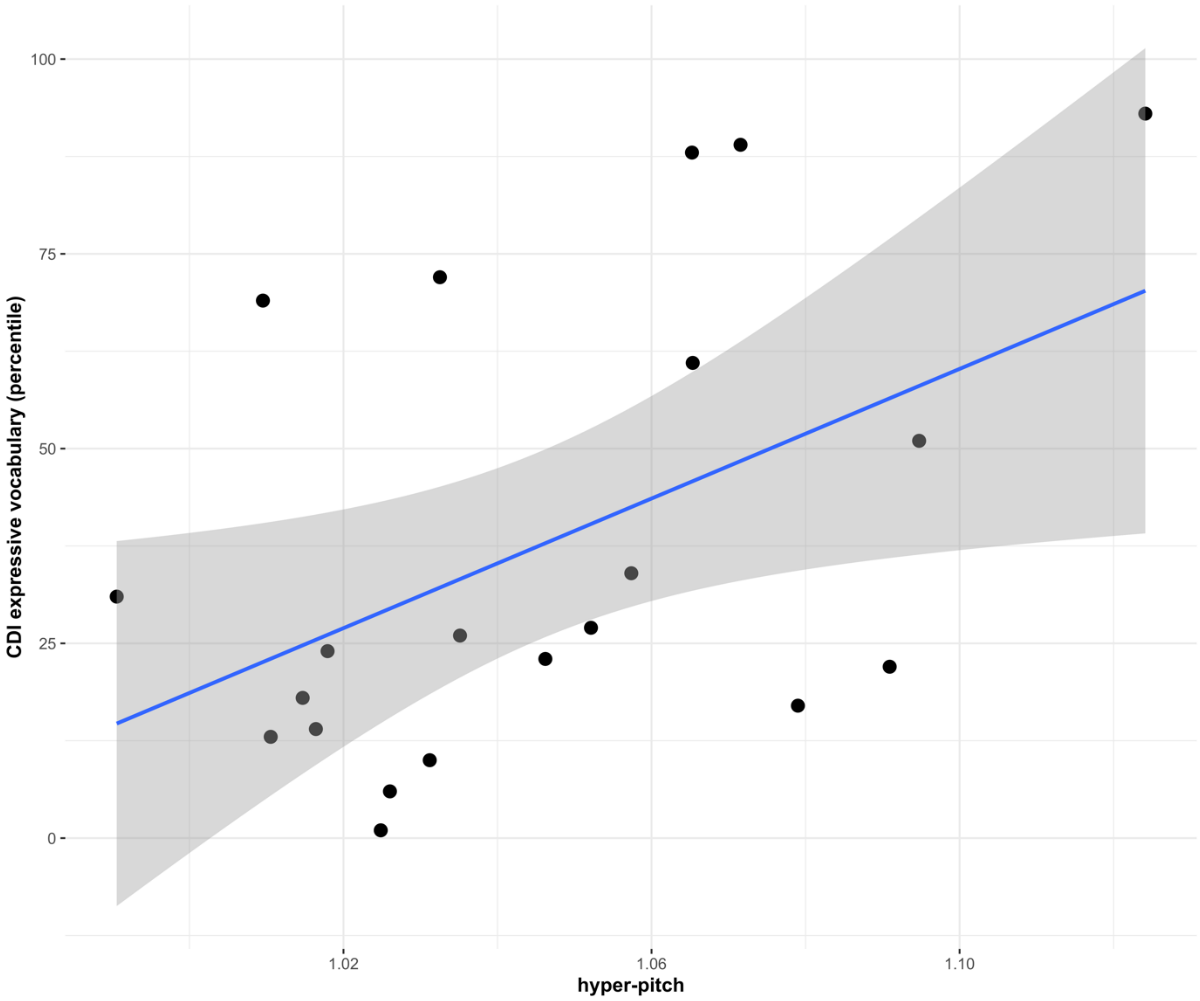


Figure 3. ***Relationship between parents' hyper-pitch and toddlers' vocabulary. Note that the figure visualises the regression line, with the shaded area depicting 95% confidence intervals. Hyper-pitch is the within-parent difference ratio of average pitch, in semitones, in IDS vs ADS.***

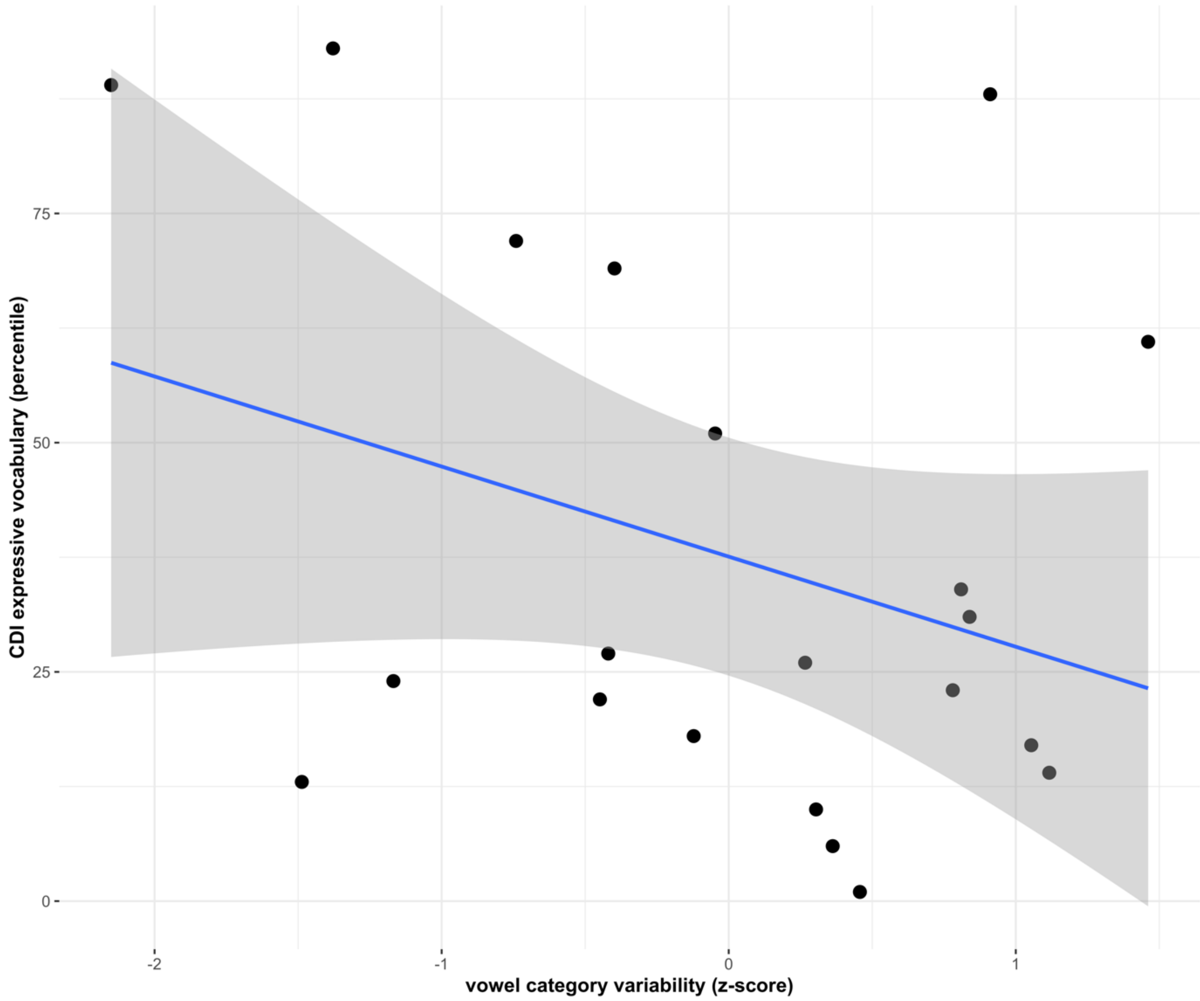


Figure 4. ***Relationship between parents' vowel category variability in IDS and toddlers' vocabulary. Note that the figure visualises the regression line, with the shaded area depicting 95% confidence intervals. The X-axis represents the z-scaled (within mothers and fathers) category variability.***

Discussion

The current study aimed to expand the knowledge about IDS in understudied languages and its potentially facilitating role in early language development. To achieve these aims, we undertook three steps: (1) examined speech of Norwegian parents speaking a Northern Norwegian dialect to their 18-month-old toddlers by measuring traditionally reported and novel acoustic properties of IDS and their differences with respect to ADS; (2) assessed the role of within-parent adaptation between IDS and ADS in predicting toddlers' expressive vocabulary, and, finally, (3) assessed the role of acoustic properties in IDS itself, in predicting toddlers' expressive vocabulary. Both IDS and ADS were elicited via a storybook reading, to control for within and between-parent differences in linguistic context that can affect speech production.

Overall, the results of the current study, in Norwegian parents to 18-month-old toddlers, supported the first hypothesis on speech 'adaptation' in IDS, as compared to ADS, providing further evidence to the growing body of research indicating that the speech register we use when interacting with young children has unique features, also in a language and a dialect that uses pitch and duration as lexical cues. Parents in our sample had higher mean phrasal pitch, wider phrasal pitch range, and longer vowel durations in IDS over ADS, although the latter was only true for mothers and not fathers. These results are in line with previous studies in other Norwegian dialects (Englund & Behne, 2005, 2006; Kartushina et al., 2021); yet, the gender differences in vowel duration suggest that fathers might be more restrained in IDS than mothers, which goes against the hypotheses that fathers' more energetic interaction style, as compared to mothers, is also manifested in IDS acoustics (Benders et al., 2021). Still, fathers in our study increased their pitch range in IDS, and thus the lack of vowel prolongation could also be related to our limited sample size for fathers, cross-linguistic differences and/or task demands, that is, a storybook reading. Further, parents expanded their vowel space area in IDS more than in ADS, both when examining the corner vowels that are typically reported in the literature (/iː/, /**ɑ**ː/, /uː/), the corner vowels particular to the Norwegian language (/iː/, /**æ**ː/, /uː/), and the full vowel space covering all border vowels in Norwegian (/ɑː/, /eː/, /iː/, /uː/, /ʉː/, /æː/, /ɔː/). This result is consistent with the studies in English (Kalashnikova & Burnham, 2018), Russian and Swedish (Kuhl et al., 1997; Marklund & Gustavsson, 2020), Spanish and Basque (Kalashnikova & Carreiras, 2021), as well as Eastern Norwegian (Kartushina et al., 2021), but not Central Norwegian (Englund & Behne, 2006). Apart from differences in the methodologies between the current and Englund and Behne's study, differences in the results on vowel space expansion between these two studies can be attributed either to fine-grained variations within a language (due to dialectal differences), or to differences in children's ages (0–6-month-old infants in Englund and Behne's study). However, vowel categories were more variable in IDS, suggesting that vowel space expansion did not necessarily translate into more intelligible speech. This supports previous work showing more variable underlying vowel categories in speech addressed to infants and toddlers (Cristia & Seidl, 2014; Martin et al., 2015; McMurray et al., 2013; Miyazawa et al., 2017). Furthermore, in Norwegian, such variability has been found in speech to 8–9-month-old infants (Kartushina et al., 2021), and now to 18-month-old toddlers, suggesting no changes in variability with the child's age. The 'sloppiness' of vowel production in IDS could potentially be a side effect of a larger vowel space expansion, or increased pitch variability, that impacts both F1 and F2 (McMurray et al., 2013). Finally, vowel category distinctiveness was comparable across registers, suggesting that although the vowel space was expanded, and the variability of individual vowel categories was increased in parents' IDS, the vowel type did not appear less identifiable within the participants' vowel clusters across registers. This could be due to parents taking extra care due to the rich vowel inventory of Norwegian, encompassing a total of 19 categories (nine long, nine short, plus schwa). Future work should expand on this result by assessing a bigger range of vowel tokens per participant, and preferably in other languages and dialects that have closer or more distributed mappings of their vowels in F1-F2 space.

With respect to our second hypothesis on the role of differences between IDS and ADS in early language development, our results showed that parents' hyper-pitch predicted toddlers' vocabulary, whereas the other acoustic measures included in our model did not. In other words, those parents who exaggerated their average pitch to a greater degree when reading to their toddlers (as compared to an experimenter), had toddlers with larger vocabulary sizes. Experimental studies have similarly highlighted the role of pitch, in supporting word segmentation in 9-month-old infants (Schreiner & Mani, 2017), and word learning in older toddlers (Graf Estes & Hurley, 2013). Recall, that increase in pitch has been reported as one of the few acoustic features present in the majority of the examined studies, suggesting it to be one of the most salient cues in IDS. In addition, research has shown that infants display larger preference for IDS at older ages (The ManyBabies Consortium, 2020), and this preference is suggested to be driven mainly by pitch increase (Segal & Newman, 2015). Thus, such a preference might engage parents in using higher pitch when interacting with their toddlers, as toddlers might be more responsive in return. As Norwegian uses pitch accent as both a lexically contrastive cue and a cue to mark dialects, parents' pitch increase, as shown in the current study, might also help toddlers incorporate these cues, thus scaffolding the development of their vocabulary.

Finally, with respect to our third hypothesis that addressed the role of direct acoustic infant-directed input in early language development, vowel category variability correlated negatively with toddlers' expressive vocabulary. This result suggests that input containing more precise vowels with little variability within each vowel category may provide scaffolding cues to build a richer vocabulary as reliable vowel productions would facilitate phonological discrimination and establishment of more stable phonological representations (see e.g., Bosch & Ramon-Casas, 2011; Bosch & Ramon-Casas, 2009; Cristia, 2011), facilitating, in turn, the vocabulary acquisition. Although laboratory studies have found facilitatory effects of vowel space expansion on speech processing (Peter et al., 2016; Song et al., 2010), experimental stimuli are *de-facto* less variable, and thus, compact categories might play more important role in 'real life' input, as compared to an experimental setting. Our result is in contrast with that of Hartman and colleagues (2017), who found that vowel space area in IDS, and not vowel variability, predicted vocabulary in similar aged English-learning toddlers. This discrepancy in the results could be due to cross-linguistic differences in vowel realization and variability and/or to differences in the number of analysed vowels; note that Hartman and colleagues examined the three corner vowels only, which might not have captured parents' full vowel inventory, as attempted in the current study with all Norwegian long vowels.

Crucially, our study demonstrates that the properties of IDS that relate to language outcomes might depend on whether the IDS is operationalised as the acoustic input directed towards the child, or as a within-parent perceptual adaptation when addressing their speech to a child as compared to an adult, respectively. It might be that hyper-pitch as a predictor of vocabulary does not reflect benefits of the acoustic signal *per se*, but rather parents' investment in capturing the attention of their toddlers, and thus such hyper-measures might be better thought of as an index of engagement and parenting style, rather than barely an acoustic booster. Although we did not find any relationship between parents' attitudes towards book reading and the quality of the linguistic input in early childhood and their degree of hyper-pitch, these were exploratory analyses and were not necessarily suited to untangle such a relationship. On the contrary, parents' precision in vowel production when interacting with their children, regardless of the differences with the ADS, correlated with their toddler's vocabulary size. Within this framework it seems more plausible to suggest benefits directly related to the acoustic signal of speech itself. Yet, we note that both of these findings are purely correlational. We need to acknowledge that third variables, such as the time parents spent with the child, or the SES – lacking diversity in our sample – might be mediating these relationships. It has also been suggested that linguistic input has the best function when it is tailored to and matches the linguistic level of the child (Rowe & Snow, 2020). Precisely, recent studies suggest that parents are experts in tuning their speech to their toddlers' needs, both lexically (Leung et al., 2021), but also acoustically (Han et al., 2020, 2021). As such, vocabulary size and parent input might be bi-directional in nature: Toddlers with a richer vocabulary (as opposed to poor) may encourage parents to increase their engagement during storybook reading more (i.e., with hyper-pitch), which, in turn, can lead to clearer (engaging, scaffolding) input to the child.

The current study has several limitations that can be addressed in future research. First, given that we did not target mothers and fathers specifically, but asked the primary caregiver to come to the lab, fathers were underrepresented in our sample, not allowing us to evaluate parent gender differences in IDS more systematically, which have been illustrated elsewhere (Benders et al., 2021). Given that Norway is a highly egalitarian society, where fathers, through the social policy, are promoted as equally important and invested caregivers with the same number of weeks of parental leave as mothers (Brandth & Kvande, 2020), this should be further investigated. Second, we used parent-reported vocabulary as our outcome measure, and although the CDI has shown to be convergent with direct child-based measures of word comprehension (Lo et al., 2021), there is a need to connect properties of IDS with direct language measures in children. Finally, the current study only captured a particular moment in time, and as parents' IDS might change across development (Narayan &

McDermott, 2016), and thus exercise varying influence on language outcomes (McMurray et al., 2013; Rowe & Snow, 2020), longitudinal studies that depict these trajectories would provide stronger evidence of such trajectories over time.

In sum, the current study provides evidence that IDS to 18-month-old Norwegian toddlers follows the same prosodic characteristics as typically reported in the literature for other languages, including increased pitch, pitch range, vowel duration (for mothers), as well as vowel space expansion, although previously reported absent in Norwegian parents to 6-month-olds (Englund, 2018). Yet, additional analyses revealed that parents' vowel categories were more variable in IDS than ADS, in line with previous research, providing evidence that parental vowel categories in IDS are less consistent and more overlapping than in ADS. Furthermore, our study indicates that hyper-pitch as well as low vowel category variability in IDS were positively associated with toddlers' vocabulary. Although the direction and the cause of the effects cannot be asserted with our design, this suggests that parents' increase in pitch when interacting with their child and their consistency in vowel production may facilitate early word learning.

References

Audibert, N., & Falk, S. (2018). Vowel space and f0 characteristics of infant-directed singing and speech. *Proceedings of the International Conference on Speech Prosody*, 153–157. https://doi.org/10.21437/SpeechProsody.2018-31

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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

Benders, T., StGeorge, J., & Fletcher, R. (2021). Infant-directed speech by Dutch fathers: Increased pitch variability within and across utterances. *Language Learning and Development*, *0*(0), 1–34. https://doi.org/10.1080/15475441.2021.1876698

Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer [Computer program]*. http://www.praat.org/

Bosch, L., & Ramon-Casas, M. (2009). Phonetic variability in bilinguals' acquisition of native-vowel category contrasts. *The Journal of the Acoustical Society of America*, *125*(4), 2770–2770. https://doi.org/10.1121/1.4784720

Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics, 39*(4), 514–526. https://doi.org/10.1016/j.wocn.2011.02.001

Brandth, B., & Kvande, E. (2020). *Designing Parental Leave Policy: The Norway Model and the Changing Face of Fatherhood*. Policy Press.

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. https://doi.org/10.1080/15248372.2013.833923

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 & Hearing Research, 58*(2), 241–253. https://doi.org/10.1044/2015_JSLHR-S-13-0205

Byers-Heinlein, K., Tsui, A. S. M., Bergmann, C., Black, A. K., Brown, A., Carbajal, M. J., Durrant, S., Fennell, C. T., Fiévet, A.-C., Frank, M. C., Gampe, A., Gervain, J., Gonzalez-Gomez, N., Hamlin, J. K., Havron, N., Hernik, M., Kerr, S., Killam, H., Klassen, K., … Wermelinger, S. (2021). A multilab study of bilingual infants: Exploring the preference for infant-directed speech. *Advances in Methods and Practices in Psychological Science, 4*(1), 2515245920974622. https://doi.org/10.1177/2515245920974622

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

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*(5), 1584–1595. https://doi.org/10.1111/j.1467-8624.1990.tb02885.x

Cribari-Neto, F., & Zeileis, A. (2010). Beta Regression in R. *Journal of Statistical Software, 34*, 1–24. https://doi.org/10.18637/jss.v034.i02

Cristia, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *Journal of the Acoustical Society of America, 129*(5), 3271–3280. https://doi.org/10.1121/1.3562562

Cristia, A. (2013). Input to language: the phonetics and perception of infant-directed speech. *Language and Linguistics Compass, 7*(3), 157–170. https://doi.org/10.1111/lnc3.12015

Cristia, A. (2022). A systematic review suggests marked differences in the prevalence of infant-directed vocalization across groups of populations. *Developmental Science,* e13265. https://doi.org/10.1111/desc.13265

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

Englund, K. T. (2018). Hypoarticulation in infant-directed speech. *Applied Psycholinguistics, 39*(01), 67–87. https://doi.org/10.1017/s0142716417000480

Englund, K. T., & Behne, D. (2005). Infant directed speech in natural interaction—Norwegian vowel quantity and quality. *Journal of Psycholinguistic Research, 34*(3), 259–280.

Englund, K. T., & Behne, D. (2006). Changes in infant directed speech in the first six months. *Infant and Child Development: An International Journal of Research and Practice, 15*(2), 139–160.

Farran, L. K., Lee, C.-C., Yoo, H., & Oller, D. K. (2016). Cross-cultural register differences in infant-directed speech: An initial study. *PloS One, 11*(3), e0151518. https://doi.org/10.1371/journal.pone.0151518

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

Fernald, A., Taeschner, T., Dunn, J., Papoušek, M., 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,* 477–501.

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

Fox, J., & Weisberg, S. (2018). *An R Companion to Applied Regression.* SAGE Publications.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data*. *Journal of Child Language, 44*(3), 677–694. https://doi.org/10.1017/S0305000916000209

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*, 104890. https://doi.org/10.1016/j.bandl.2020.104890

Garnier, M., Ménard, L., & Alexandre, B. (2018). Hyper-articulation in Lombard speech: An active communicative strategy to enhance visible speech cues? *The Journal of the Acoustical Society of America, 144*(2), 1059–1074. https://doi.org/10.1121/1.5051321

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

Graf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy, 18*(5), 797–824. https://doi.org/10.1111/infa.12006

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

Han, M., de Jong, N. H., & Kager, R. (2021). Language specificity of infant-directed speech: speaking rate and word position in word-learning contexts. *Language Learning and Development, 17*(3), 221–240. https://doi.org/10.1080/15475441.2020.1855182

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

Hembacher, E., & Frank, M. C. (2020). The early parenting attitudes questionnaire: Measuring intuitive theories of parenting and child development. *Collabra: Psychology, 6*(1). https://doi.org/10.1525/collabra.190

Hirst, D. (2012). *Analyse_tier.praat* [Praat script]. https://www.mail-archive.com/praat-users@yahoogroups.co.uk/msg00061.html

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*(8), 170306. https://doi.org/10.1098/rsos.170306

Kalashnikova, M., & Carreiras, M. (2021). Input quality and speech perception development in bilingual infants' first year of life. *Child Development*, *n/a*(n/a). https://doi.org/10.1111/cdev.13686

Kartushina, N., & Frauenfelder, U. H. (2014). On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology*, *5*, 1246. https://doi.org/10.3389/fpsyg.2014.01246

Kartushina, N., Mani, N., Aktan-Erciyes, A., Alaslani, K., Aldrich, N. J., Almohammadi, A., ..., & Mayor, J. (2022). COVID-19 first lockdown as a window into language acquisition: associations between caregiver-child activities and vocabulary gains. *Language Development Research*, *2*(1), 1–36. https://doi.org/10.34842/abym-xv34

Kartushina, N., Robbestad, S., & Mayor, J. (2021, April 7). *The Role of Parental Speech in Infant Language Development: Insights from 8-9-month-old Norwegian Infants*. Society for Research in Child Development 2021 Virtual Biennial Meeting. https://www.srcd.org/event/srcd-2021-biennial-meeting

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

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.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.
Leung, A., Tunkel, A., & Yurovsky, D. (2021). Parents fine-tune their speech to children's vocabulary knowledge. *Psychological Science*, *32*(7), 975–984. https://doi.org/10.1177/0956797621993104

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

Lo, C. H., Rosslund, A., Chai, J. H., Mayor, J., & Kartushina, N. (2021). Tablet assessment of word comprehension reveals coarse word representations in 18–20-month-old toddlers. *Infancy*, *26*(4), 596–616. https://doi.org/10.1111/infa.12401

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

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

Marklund, E., & Gustavsson, L. (2020). The dynamics of vowel hypo- and hyperarticulation in Swedish infant-directed speech to 12-month-olds. *Frontiers in Communication, 5,* 523768. https://doi.org/10.3389/fcomm.2020.523768

Marklund, E., Marklund, U., & Gustavsson, L. (2021). An association between phonetic complexity of infant vocalizations and parent vowel hyperarticulation. *Frontiers in Psychology, 12,* 2873. https://doi.org/10.3389/fpsyg.2021.693866

Monsen, R. B., & Engebretson, A. M. (1983). The accuracy of formant frequency measurements: A comparison of spectrographic analysis and linear prediction. *Journal of Speech, Language, and Hearing Research, 26*(1), 89-97. https://doi.org/10.1044/jshr.2601.89

Mæhlum, B., & Røyneland, U. (2012). *Det norske dialektlandskapet*. Cappelen Damm Akademisk.

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

McClay, E. K., Cebioglu, S., Broesch, T., & Yeung, H. H. (2021). Rethinking the phonetics of baby-talk: Differences across Canada and Vanuatu in the articulation of mothers' speech to infants. *Developmental Science, n/a*(n/a), e13180. https://doi.org/10.1111/desc.13180

McCloy, D. (2016). *PhonR: tools for phoneticians and phonologists* (1.0-7) [Computer software]. https://cran.r-project.org/web/packages/phonR/phonR.pdf

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

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

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. https://doi.org/10.1080/15475441.2013.802962

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. www.R-project.org

Rowe, M. L., & Snow, C. E. (2020). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. *Journal of Child Language, 47*(1), 5–21. https://doi.org/10.1017/S0305000919000655

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–e78103. https://doi.org/10.1371/journal.pone.0078103

Schreiner, M. S., & Mani, N. (2017). Listen up! Developmental differences in the impact of IDS on speech segmentation. *Cognition, 160*, 98–102. https://doi.org/10.1016/j.cognition.2016.12.003

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. https://doi.org/10.1111/infa.12077

Simonsen, H. G., Kristoffersen, K. E., Bleses, D., Wehberg, S., & Jørgensen, R. N. (2014). The Norwegian communicative development inventories: Reliability, main developmental trends and gender differences. *First Language, 34*(1), 3–23. https://doi.org/10.1177/0142723713510997

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

Steinlen, A., & Bohn, O. (1999). Acoustic studies comparing Danish vowels, British English vowels, and Danish-accented British English vowels. *The Journal of the*

Acoustical Society of America, 105(2), 1097–1097. https://doi.org/10.1121/1.425143

Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science, 20*(6), e12456. https://doi.org/10.1111/desc.12456

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

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

Wang, Y., Seidl, A., & Cristia, A. (2015). Acoustic-phonetic differences between infant- and adult-directed speech: The role of stress and utterance position. *Journal of Child Language, 42*(4), 821–842. https://doi.org/10.1017/S0305000914000439

Wichmann, S. (2019). How to distinguish languages and dialects. *Computational Linguistics, 45*(4), 823–831. https://doi.org/10.1162/coli_a_00366

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

Xu Rattanasone, N., 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

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution, 1*(1), 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

Data, Code and Materials Availability Statement

Data, code and materials are available at the Open Science Framework project's page, at the following permanent link: https://osf.io/7st6w/.

Ethics Statement

The current study was conducted according to the guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or a guardian for a child before any assessment or data collection. The study has been approved by the Norwegian Centre for Research Data (NSD, ref. 56312), and the local ethical committee at the Department of Psychology, University of Oslo.

Authorship and Contributorship Statement

AR, JM, GÓ and NK conceptualised the study. JM, GÓ and NK supervised the study. AR analysed the data and wrote the first draft of the manuscript. AR, JM, GÓ and NK interpreted the results and revised the manuscript. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

We thank Mathea Sandholm, Hanne Akseth Ulriksen, Madelene Halvari Niska, and Annie-Justicia Karlsson for their help with data collection, and Roger Mundry for help with the vowel category distinctiveness measure. We would like to thank reviewers for their insightful comments and suggestions. AR and NK were supported by the Research Council of Norway through project number 301625, and its Centres of Excellence funding scheme [project number 223265].

Appendices

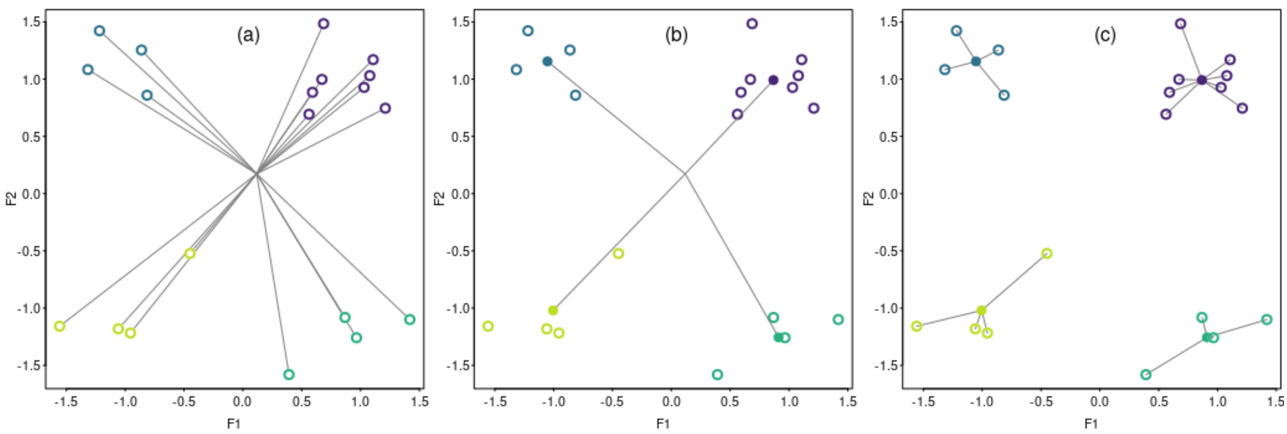
Appendix 1. ***Overview of words in the storybook for each target vowel (in bold)***

/iː/	/yː/	/eː/	/øː/	/æː/	/ʉː/	/uː/	/ɔː/	/ɑː/
b**i**l (car)	l**y**s (light)	s**e** (look)	br**ø**d (bread)	d**e**r (there)	l**u**e (hat)	b**o**k (book)	s**o**ve (sleep)	ban**a**n (banana)
gr**i**s (pig)	fl**y** (airplane)	skj**e** (spoon)	sn**ø** (snow)	h**e**r (here)	p**u**te (pillow)	sk**o** (shoe)	t**o**g (train)	b**a**de (bath)
sp**i**se (eat)	d**y**ne (duvet)	m**e**r (more)	d**ø**r (door)	v**æ**re (be)	k**u** (cow)	f**o**t (foot)	h**å**r (hair)	k**a**ke (cake)
sk**i**ve (slice)	d**y**r (animal)	n**e**se (nose)	bj**ø**rn (bear)	b**æ**re (carry)	m**u**s (mouse)	s**o**l (sun)	m**å**ne (moon)	m**a**ge (belly)
v**i** (we)	n**y** (new)	l**e**se (read)	l**ø**pe (run)	skj**æ**re (cut)	f**u**gl (bird)	hall**o** (hello)	g**å** (go)	br**a** (good)

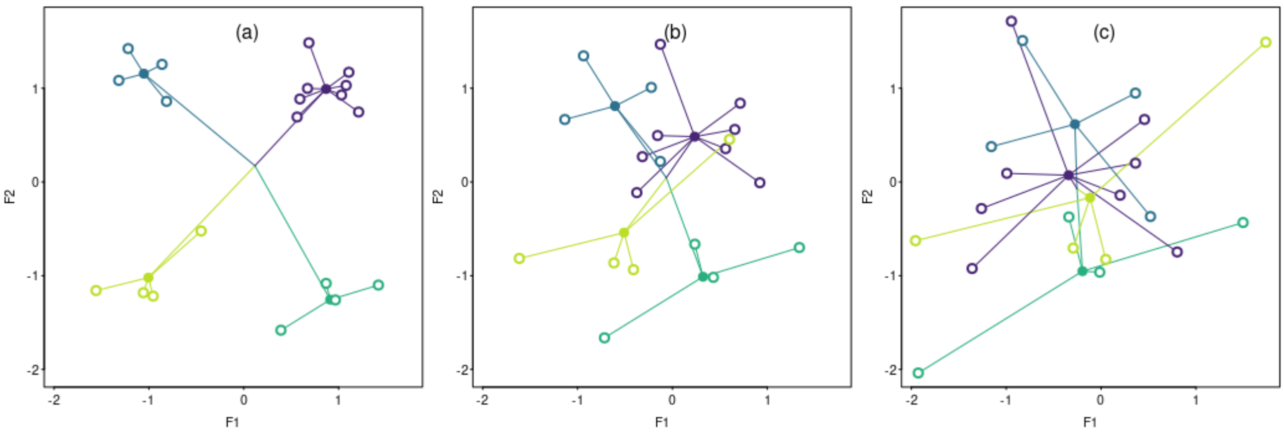
Appendix 2. *Comparison of two 'born-early' dyads to the rest of the sample*

Variable	Not 'born early' (n = 19)	'Born early' (n = 2)	Full sample (n = 21)
Cat_dist.mean.IDS			
Mean (SD)	0.885 (0.0500)	0.903 (0.0203)	0.887 (0.0480)
Median [Min, Max]	0.900 [0.745, 0.952]	0.903 [0.889, 0.918]	0.900 [0.745, 0.952]
Cat_var.mean.IDS			
Mean (SD)	317000 (132000)	249000 (78200)	311000 (128000)
Median [Min, Max]	279000 [109000, 551000]	249000 [194000, 304000]	279000 [109000, 551000]
Vowel_space_full_IDS			
Mean (SD)	433 (106)	521 (193)	441 (113)
Median [Min, Max]	428 [223, 615]	521 [385, 658]	428 [223, 658]
Vowel_space_æ_IDS			
Mean (SD)	433 (106)	521 (193)	441 (113)
Median [Min, Max]	428 [223, 615]	521 [385, 658]	428 [223, 658]
Duration.mean.IDS			
Mean (SD)	132 (21.8)	150 (23.7)	134 (22.0)
Median [Min, Max]	128 [96.0, 187]	150 [133, 167]	128 [96.0, 187]
Pitch_range.mean.IDS			
Mean (SD)	14.9 (3.41)	12.4 (4.55)	14.6 (3.47)
Median [Min, Max]	14.7 [6.41, 22.2]	12.4 [9.21, 15.7]	14.7 [6.41, 22.2]
Pitch.mean.IDS			
Mean (SD)	53.8 (6.38)	54.7 (2.10)	53.9 (6.08)
Median [Min, Max]	56.5 [40.7, 61.4]	54.7 [53.2, 56.2]	56.2 [40.7, 61.4]
Cat_dist_effort			
Mean (SD)	1.00 (0.0569)	1.00 (0.0254)	1.00 (0.0543)
Median [Min, Max]	1.01 [0.859, 1.10]	1.00 [0.985, 1.02]	1.01 [0.859, 1.10]
Cat_var_effort			
Mean (SD)	1.20 (0.373)	1.28 (0.00276)	1.21 (0.355)
Median [Min, Max]	1.10 [0.641, 2.08]	1.28 [1.28, 1.28]	1.15 [0.641, 2.08]
Vowel_space_full_effort			
Mean (SD)	1.15 (0.269)	1.53 (0.760)	1.19 (0.327)
Median [Min, Max]	1.13 [0.667, 1.90]	1.53 [0.992, 2.07]	1.13 [0.667, 2.07]
Vowel_space_æ_effort			
Mean (SD)	1.15 (0.268)	1.23 (0.146)	1.15 (0.257)
Median [Min, Max]	1.07 [0.758, 1.73]	1.23 [1.13, 1.34]	1.09 [0.758, 1.73]
Duration_effort			
Mean (SD)	1.18 (0.162)	1.07 (0.195)	1.17 (0.163)
Median [Min, Max]	1.15 [0.945, 1.65]	1.07 [0.935, 1.21]	1.15 [0.935, 1.65]
Pitch_range_effort			
Mean (SD)	1.12 (0.239)	1.16 (0.315)	1.13 (0.238)
Median [Min, Max]	1.17 [0.655, 1.47]	1.16 [0.935, 1.38]	1.17 [0.655, 1.47]
Pitch_effort			
Mean (SD)	1.05 (0.0341)	1.02 (0.00229)	1.05 (0.0338)
Median [Min, Max]	1.05 [0.991, 1.12]	1.02 [1.01, 1.02]	1.04 [0.991, 1.12]
CDI_prod_percentile			
Mean (SD)	39.3 (30.3)	21.0 (4.24)	37.6 (29.3)
Median [Min, Max]	27.0 [1.00, 93.0]	21.0 [18.0, 24.0]	26.0 [1.00, 93.0]

Appendix 3A. ***Illustration of the method to determine cluster distinctiveness. The total sum of squares (SStot) is the sum of the squared distances of the individual vocalizations from the overall centroid (a). The between cluster sum of squares (SSbetween) is sum of the squared distances of the per-vowel centroids times the number vocalizations per vowel (b). The within cluster sum of squares is the sum of the squared distances of the individual vocalizations from the respective vowel's centroid (c). Each vowel is depicted by a specific color, individual vocalizations by open dots, and vowel centroids by filled dots. Note the SStot = SSbetween + SSwithin. In case of the example SStot = 43.544, SSbetween = 40.609, SSwithin = 2.936, and cluster distinctiveness = 0.933.***



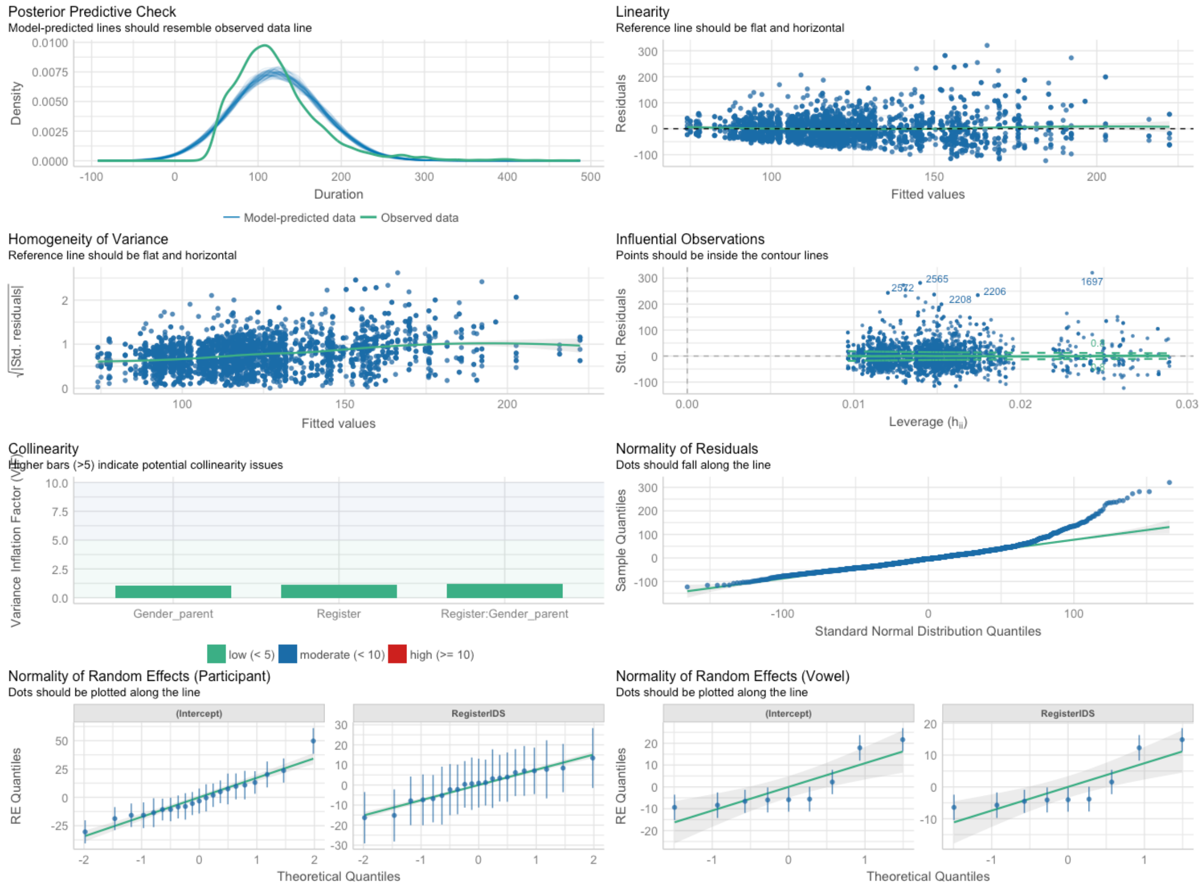
Appendix 3B. ***Illustration of various values of cluster distinctiveness. Each vowel is depicted by a different color, open dots show the individual utterances, and filled dots the clusters' centroids. The total variance explained by vowel type ('cluster distinctiveness') is 0.93 in (a), 0.53 in (b), and 0.14 in (c).***



Appendix 4. ***sessionInfo() output providing R libraries and their versions that were used in the analyses***

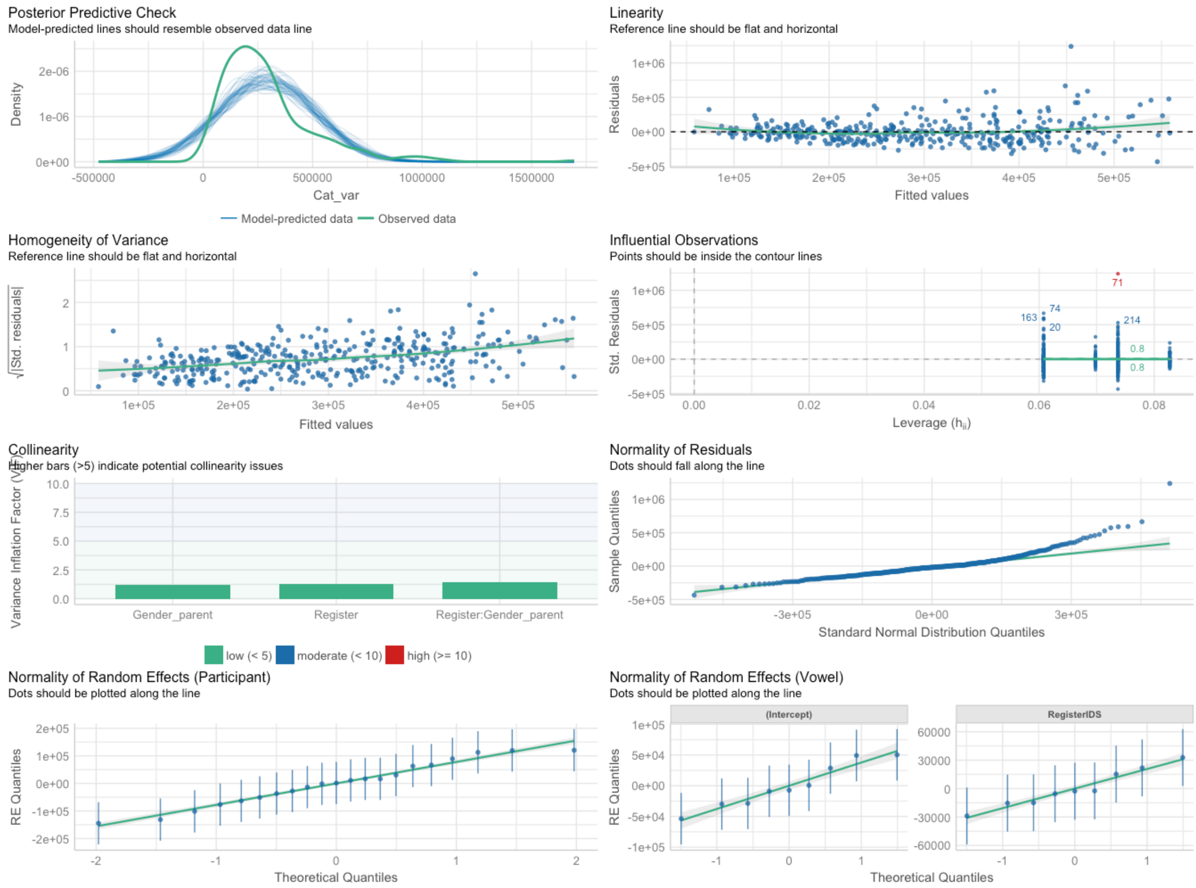
Package	Loaded version	Date
betareg	3.1-4	2021-02-09
car	3.0-12	2021-11-06
carData	3.0-4	2020-05-22
doBy	4.6.11	2021-07-13
dplyr	1.0.7	2021-06-18
effsize	0.8.1	2020-10-05
emmeans	1.7.1-1	2021-11-29
emuR	2.3.0	2021-06-11
factoextra	1.0.7	2020-04-01
forcats	0.5.1	2021-01-27
ggplot2	3.3.5	2021-06-25
ggpubr	0.4.0	2020-06-27
ggstatsplot	0.9.0	2021-10-19
knitr	1.36	2021-09-29
lme4	1.1-27.1	2021-06-22
lmerTest	3.1-3	2020-10-23
lsmeans	2.30-0	2018-11-02
Matrix	1.3-4	2021-06-01
patchwork	1.1.1	2020-12-17
performance	0.8.0	2021-10-01
phonR	1.0-7	2016-08-25
purrr	0.3.4	2020-04-17
qqplotr	0.0.5	2021-04-23
rcompanion	2.4.6	2021-11-21
readr	2.1.1	2021-11-30
readxl	1.3.1	2019-03-13
shinyBS	0.61	2015-03-31
soundgen	2.5.0	2021-11-21
stringr	1.4.0	2019-02-10
table1	1.4.2	2021-06-06
tibble	3.1.6	2021-11-07
tidyr	1.1.4	2021-09-27
tidyverse	1.3.1	2021-04-15
viridis	0.6.2	2021-10-13
viridisLite	0.4.0	2021-04-13
vowels	1.2-2	2018-03-05

Appendix 5A. ***Model diagnostics of vowel duration pre log-transformation***



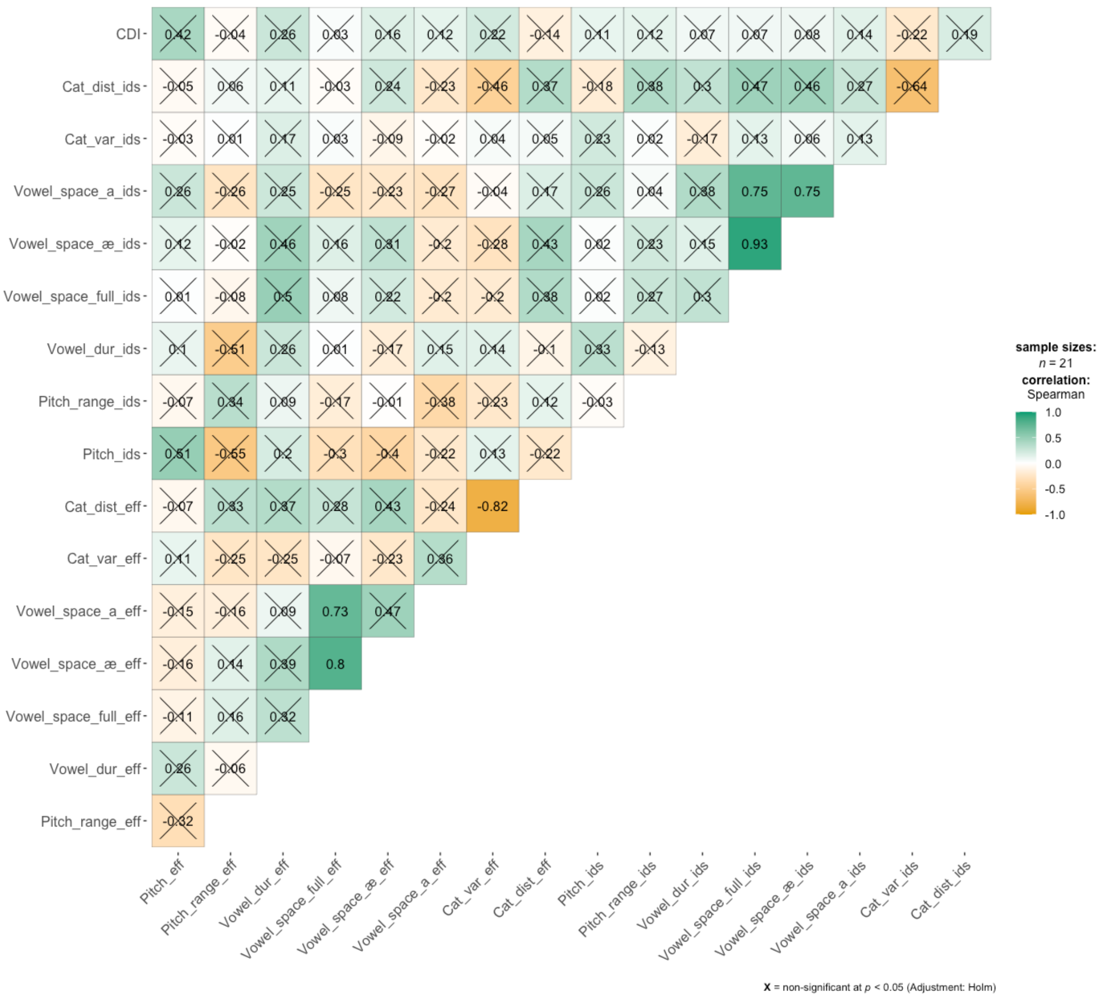
Appendix 5B. *Model diagnostics of vowel duration post log-transformation*

Appendix 6A. *Model diagnostics of vowel category variability pre log-transformation*

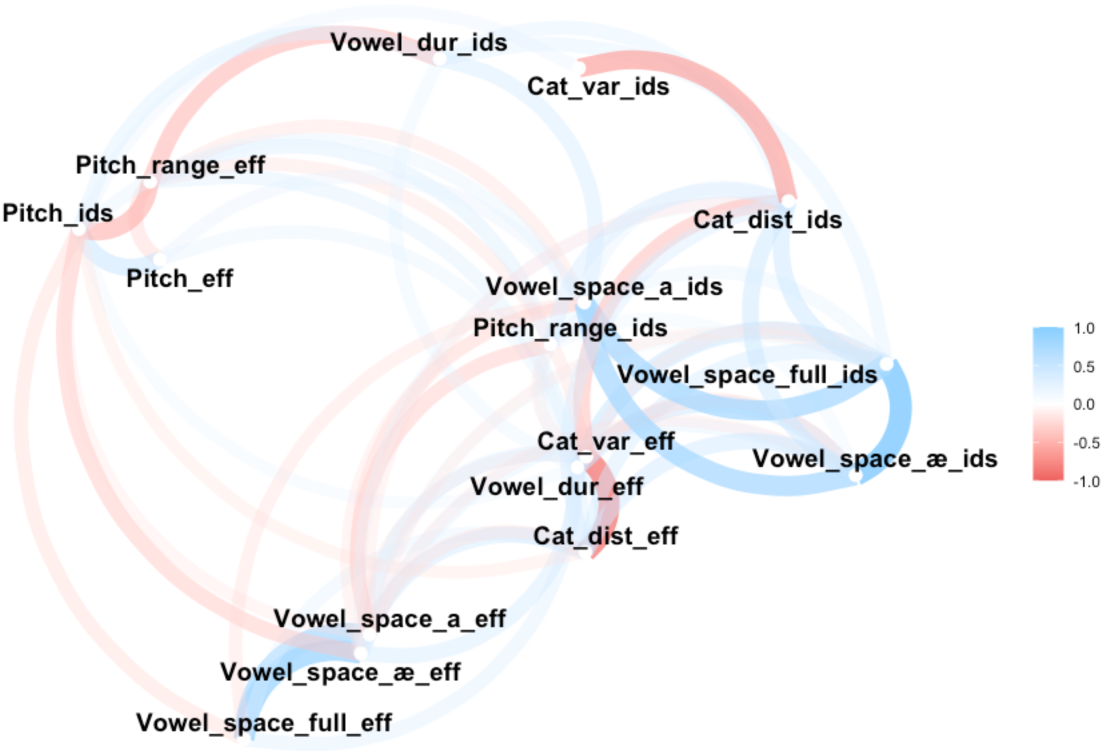


Appendix 6B. ***Model diagnostics of vowel category variability post log-transformation***

Appendix 7A. ***Spearman correlation matrix of difference (_eff) and IDS-input (_ids) acoustic measures***



Appendix 7B. ***Spearman correlation network of difference (_eff) and IDS-input (_ids) acoustic measures*****. Note that only correlations stronger than +/- .20 are displayed**



License