

BRIEF RESEARCH REPORT

Voice quality modifications in Hungarian infant-directed speech: A longitudinal acoustic study

Anna Kohári , Uwe D. Reichel and Katalin Mády

Phonetics Research Group, ELTE Research Centre for Linguistics, Budapest, Hungary

Corresponding author: Anna Kohári; Email: kohari.anna@nytud.hun-ren.hu

(Received 29 December 2024; revised 02 July 2025; accepted 22 December 2025)

Abstract

Adults exhibit different acoustic characteristics in infant-directed speech (IDS) compared to adult-directed speech (ADS). We investigate differences in voice quality between IDS and ADS in sentences read aloud by Hungarian mothers, using longitudinal data gathered at various child ages (4, 8, 18 months). Vowels in IDS are found to be breathier than those in ADS, regardless of the infant's age. Possible motivations for this difference may include emotional expressions, as breathiness relates to positive emotions, and speech entrainment, since the speech of children is breathier than that of adults.

Keywords: infant-directed speech; voice quality; breathy voice; longitudinal analysis; acoustic phonetics

Absztrakt

A dajkanyelv több akusztikai tulajdonsága eltérhet a felnőttekhez szóló beszéd sajátosságaitól. Magyar anyanyelvű anyák felolvasott mondatain vizsgáltuk a dajkanyelv zöngemínőségbeli sajátosságait longitudinálisan a gyerek különböző életkoraiban (4, 8 és 18 hónapos korában). Eredményeink szerint a dajkanyelv magánhangzói leheletesebbnek bizonyultak a felnőttekhez szóló beszédhez képest a gyerek életkorától függetlenül. A regiszterek közti zöngemínőség-különbséget motiválhatja az érzelmek kifejezése, mivel a leheletes zöngé a pozitív érzelmek kifejezője lehet, továbbá a beszédalkalmazkodás is magyarázhatja az eltérést, hiszen a gyerekeknek leheletesebb a zöngéje a felnőtteknél.

1. Introduction

When talking to infants, adults adjust to the situation and modify numerous properties of their speech, including its syntactic, semantic and acoustic characteristics. This special register is referred to as infant-directed speech (IDS) and is typically different from the way adults talk to each other (adult-directed speech, ADS). Higher fundamental frequency, lower speech and articulation rate, simplified grammar, and more frequent repetitions compared to ADS are all considered to be common features of IDS (Genovese et al., 2020; Harmati-Pap et al., 2024; Kuhl et al., 1997; Saint-Georges et al., 2013).

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Although voice quality is not widely regarded as a typical characteristic of IDS, several recent studies have indicated that speakers modify this feature as well when switching from ADS to IDS (Cheng et al., 2024; McClay et al., 2022; Miyazawa et al., 2017). The vibration of the vocal folds can lead to multiple phonation types, depending on the size of the glottal opening (Garellek, 2019; Ladefoged, 1973). The different phonation types form a continuum. A typical breathy voice is produced when voicing occurs with less vocal fold approximation. In the prototypical creaky voice, phonation happens with more approximation. The modal voice is somewhere between these two extremes. Vocal fold vibration yields a periodic waveform in modal voice, whereas in breathy and creaky voices, aperiodic noise also emerges (Garellek, 2019). Three recent studies have found that Canadian English, Czech, and Japanese native-speaker mothers produce a breathier voice in IDS than in ADS (Cheng et al., 2024; Chládková et al., 2019; Miyazawa et al., 2017). These works applied widely used acoustic measures of vocal fold approximation or the periodicity in the acoustic signal relative to its noisiness for the analysis of voice quality. Cheng et al. (2024) have found that Canadian English-speaking mothers use a breathier voice in IDS than in ADS. Their results were based on a re-analysis of material from an earlier study that had reported entirely different conclusions (McClay et al., 2022). McClay et al. (2022) did not report any difference between the two registers, but Cheng et al. (2024) found such differences in their later, more detailed analysis. In McClay et al. (2022), the authors also reported on the speech characteristics of mothers from Vanuatu, whose IDS was found to be less breathy than their ADS, implying an opposite trend than in other investigated languages. However, the Vanuatu results were not re-analysed in their later work that included more detailed investigations.

The question arises of what purpose a breathy voice may serve in IDS. It has been suggested that speakers use certain acoustic features of IDS (e.g., increased fundamental and formant frequencies) to appear less aggressive or threatening to infants since these characteristics generally coincide with smaller body sizes (Kalashnikova et al., 2017). Moreover, it has also been proposed that the properties of IDS may serve the purpose of signalling positive emotions and approachability (Benders, 2013; Hilton et al., 2022; Mády et al., 2022). Generally, children tend to show a preference for IDS over ADS, and their attention is also increased when they listen to IDS (Háden et al., 2020; ManyBabies Consortium, 2020; Spinelli et al., 2017). The acoustic properties of this register may help children recognize that they are the addressees of the speech, which may contribute to the language learning process (Senju & Csibra, 2008). These features may also support the child's speech perception, yielding faster and more effective language development (Rosslund et al., 2022; Spinelli et al., 2017). As for voice quality, it is known that small children tend to produce breathier voices than older ones and adults (Kent et al., 2021; Zhang, 2021). Therefore, using a breathier voice in IDS may be motivated by trying to adapt to the child's assumed or actual speech production. It has also been shown that happy speech involving a breathy voice attracts children's attention even more than emotional speech with a modal voice (Kao et al., 2022). This may be related to the fact that positive emotions tend to be associated with a breathy voice, as reported by studies investigating English and Hungarian speech (Anikin, 2020; Bartók, 2018).

The acoustic properties of IDS may change with the age of the child. Certain features of this register, such as the lower articulation rate or the higher fundamental frequency, approach the values of ADS as the child ages (Kalashnikova & Burnham, 2018; Narayan & McDermott, 2016). Meanwhile, some other IDS features remain largely unchanged throughout the child's first 2 years, for example, the larger vowel space (Cox et al., 2023). To the best of our knowledge, no studies have yet investigated the development of

the longitudinal changes of voice quality in IDS across different ages of the same children. In the existing literature, the ages of the children vary widely. Japanese native speaker mothers were studied while interacting with their infants aged 18 to 24 months (Miyazawa et al., 2017), and Czech native speaker mothers were analysed in the prenatal stage, that is, during pregnancy (Chládková et al., 2019). In a study investigating Canadian English, the children's ages ranged from 6 to 22 months, and the authors found no relationship between age and voice quality (Cheng et al., 2024).

In the present study, we investigate whether and how speakers modify their voice quality in IDS compared to ADS in an understudied language, Hungarian. As in other languages, different voice quality types in Hungarian may also have special functions in prosodic phrasing. For example, the phrase-final boundaries are typically marked with a creaky voice in several languages (e.g., English, Japanese), including Hungarian (Garellek, 2022; Kawahara & Shinya, 2008; Markó, 2013). In a more recent pilot study, Kohári et al. (2024) investigated voice quality at the phrase boundaries in Hungarian IDS. The results showed that the vowels at phrase-final boundaries in IDS shifted from a creaky voice to a more modal voice. In the present study, we aimed to exclude such known systematic occurrences of creaky voice since we focused on the general characteristics of IDS. In Hungarian, it was also shown that vowels in hiatus positions and the initial positions of intonational phrases are often realized with creaky voice qualities, similarly to observations in English and other languages (Garellek, 2022; Markó, 2013). Yet, even in the absence of such prosodic or phonological conditions, creaky voice is still a frequent phenomenon in Hungarian, found in approximately 20% of the produced vowels (Grácz et al., 2017). Besides a creaky voice, a breathy voice also appears in Hungarian speech, typically when expressing positive emotions (Bartók, 2018). Although a creaky voice in Hungarian speech is relatively frequent even within phrases, mothers may use a breathier voice when talking to their infants, similarly to Canadian English, Czech, and Japanese speakers. In our analysis, we apply widely used and thoroughly validated voice quality indices: spectral tilt and periodicity measures for the sake of reliability (Cheng et al., 2024; Garellek, 2019). Furthermore, we employ a longitudinal design, studying the same speaker repeatedly at key stages of the infant's language development (Bergelson & Swingley, 2012, 2015; Tincoff & Jusczyk, 2012). We assumed that a breathy voice is stably present in IDS until the child starts to actively produce words. This assumption is supported by previous results, which indicated a clear difference between the two registers in this respect, even for mothers of older children (Miyazawa et al., 2017).

2. Methods

2.1. Participants

We analysed data from 20 first-time mothers who were native Hungarian speakers. The participants, aged 26–38 years ($M = 30.4$ years, $SD = 4.0$), reported no hearing or speech difficulties. All mothers lived in Budapest or nearby towns and had completed high school or higher education. Recruitment took place at the Birth Centre of the Military Hospital (Budapest, Hungary) during the birth of their first child. The participating children (14 boys, 6 girls) were all typically developing.

2.2. Materials and recording procedure

The participants were instructed to use an illustrated storybook to narrate a story. The book contained sentences to be read aloud, as well as images without text, about which the

participants were asked to tell a story. The purpose of the spontaneous speech part was to help the mothers produce the read sentences more naturally. The read sentences were included to ensure that the phonetic aspects of the measurements could be properly controlled. After familiarizing themselves with the storybook, participants first narrated the story to the experimenter (ADS) and then to their child (IDS). The experiments were repeated when the children were approximately 4, 8, and 18 months old (first session: $M = 4$ months and 8.3 days, $SD = 6.9$; second session: $M = 8$ months and 7.0 days, $SD = 9.7$; third session: $M = 18$ months and 7.9 days, $SD = 6.6$). The same storybook was used at all three age points. The timing was chosen to follow the basic stages of language development, as described in the literature (Bergelson & Swingley, 2012, 2015; Tincoff & Jusczyk, 2012). On the one hand, we intended to contrast the preverbal stage (4- and 8-month old) with a later one in which the children communicate actively and produce words (18 months). On the other hand, the intermediate 8-month age was selected to allow a comparison between this age, when infants have already begun to understand words, and their respective 4-month-old phase. In the latter, such comprehension is not assumed, yet infants in this stage already possess the ability to discover certain features of the speech signal. The recordings were made in the laboratory of the Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, with a head-mounted hypercardioid microphone (Beyerdynamic TG H74c) at a 44.1 kHz sampling rate and digitized at 16 bits using an M-Audio two-channel USB external sound card.

2.3. Data analysis

Ten read sentences were utilized for acoustic analysis, ensuring that the same vowels were examined in identical phonetic contexts (consonant environment and prosodic position) for both registers. The samples were first segmented automatically with MAUS (Kisler et al., 2017) and then manually corrected using Praat 6.1.08 (Boersma & Weenink, 2019). Vowel sounds that appeared at least four times in the read IDS and ADS of each speaker were selected for evaluation to enable statistical analyses. The following five Hungarian vowels were analysed in the present study: /ɒ/, /a:/, /ɛ/, /i/, and /o/ (Figure 1). As the voice quality of phrase-final syllables and the vowels in phrase-initial positions tends to differ from the modal voice in Hungarian speech (Markó, 2013), those

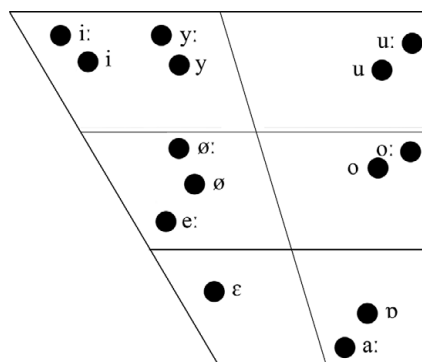


Figure 1. Schematic representation of the Hungarian vowel inventory reproduced from Markó et al. (2018).

vowels and syllables were excluded from the analysis. There were no occurrences of vowels in hiatus positions in the material. We examined 6,775 vowels in total. Acoustic parameters associated with voice quality were estimated using VoiceSauce (Shue et al., 2011). We utilized the STRAIGHT algorithm of VoiceSauce to measure F0 (Kawahara et al., 2005) with default settings. The frequencies of the vowel formants (F1, F2) were calculated using the Praat software.

For voice quality analysis, we selected three measures ($H1^*-H2^*$, $H1^*-A1^*$, and CPP) that are associated with the perception of voice quality (Garellek, 2019) and have shown differences between IDS and ADS in previous studies (Cheng et al., 2024; Miyazawa et al., 2017). F0, F1, and F2 outliers were excluded from the voice quality analysis when their values were outside the 2.5 standard deviation interval around the mean for the given speaker (cf. Garellek & Esposito, 2023). The first quantity to be determined was a measure of spectral tilt, calculated as the difference between the amplitudes corresponding to the first and second harmonics of the fundamental frequency ($H1^*-H2^*$) corrected for the effects of the vowel's formants (Iseli et al., 2007). The $H1^*-H2^*$ measure is related to the openness of the vocal folds. The more open the vocal folds are during phonation, the higher the value of this measure. In the case of a breathy voice, the vocal folds are more open, yielding higher $H1^*-H2^*$, while glottalization is characterized by low and modal voice by intermediate values (Garellek, 2019). $H1^*-A1^*$ is similar to $H1^*-H2^*$ and was earlier also considered to be a relevant measure for IDS-related phenomena (Cheng et al., 2024). Here, instead of $H2^*$ (the second harmonic of F0), the subtracted amplitude is that of the harmonic that is the closest to F1. Similarly to the previous measure, higher values of $H1^*-A1^*$ are typical for breathy voice, and lower values for creaky voice, with the modal voice in between. The Cepstral Peak Prominence (CPP) is the difference, in dB, between the dominant cepstral peak at the quefrequency $q_0 = 1/F_0$ and the value of a linear regression cepstral baseline evaluated at q_0 (Hillenbrand et al., 1994). When the vibration of the vocal folds is strongly periodic and the phonation is not noisy, the CPP is high. This is typically observed in the modal voice. When the vocal folds exhibit aperiodic vibrations with lower energy, the noise (inharmonic) components are stronger in the signal. This situation commonly occurs during the production of breathy voice and creaky voice, both of which yield lower CPP. Values of CPP have been found to show clear differences between IDS and ADS (Cheng et al., 2024).

We conducted the statistical analyses in R 4.4.0 (R Core Team, 2021). We fitted separate linear mixed-effects models for the voice quality measures ($H1^*-H2^*$, $H1^*-A1^*$, CPP), utilizing the lmerTest package (Kuznetsova et al., 2017). We followed the basic rules of model selection (Winter, 2019). The initial model was built with the fixed factors including register (IDS/ADS), infant age (4, 8, and 18 months), vowel (/b/, /a:/, /ε/, /i/, and /o/), and with all possible interactions among all these fixed factors. The maximal model included by-speaker random intercepts and slopes for all fixed effects. We simplified the random effect structure based on variance values until the model convergence was achieved. Backward stepwise model selection was performed via likelihood ratio tests using the “anova” function. The best minimal model was selected by excluding the interactions and fixed factors that did not contribute to the model. The lmerTest package (Kuznetsova et al., 2017) was applied to estimate *p*-values based on Satterthwaite approximations. Tukey's post hoc tests were carried out to account for multiple comparisons utilizing the emmeans package (Lenth et al., 2019) when needed. We estimated conditional R^2 and marginal R^2 values using the MuMIn package (Bartón, 2024).

3. Results

Descriptive statistics for all voice quality measures are reported in Table 1, separately for the registers (IDS and ADS). On average, the spectral tilt measures ($H1^*-H2^*$, $H1^*-A1^*$) seemed to be higher in IDS than in ADS, indicating the presence of the breathier voice in the former register. The CPP measure – quantifying the periodicity – showed a slight difference between the two registers, with somewhat lower values for IDS, implying that the phonation is less periodic than in ADS. This difference is also a specific characteristic of breathy voice compared to modal voice.

Based on the detailed statistical results, it can be stated with confidence that the register clearly affected the $H1^*-H2^*$ measure (Table 2). IDS typically exhibited higher values than ADS in this respect, indicating breathier phonation in IDS. However, this effect varied depending on the baby's age and the investigated vowel; therefore, we performed a post hoc test. For 4 and 18 months of the infants' age, the IDS register had higher $H1^*-H2^*$ than ADS (4 months: $b = -1.16$, $SE = 0.29$, $df = 72.50$, $t = -4.06$, $p < 0.001$; 18 months: $b = -0.71$, $SE = 0.29$, $df = 73.30$, $t = -2.50$, $p = 0.015$). When the infants were 8 months old, the same trend still appeared, but the register had only a marginal effect on this voice quality measure ($b = -0.45$, $SE = 0.29$, $df = 71.20$, $t = -1.58$, $p = 0.120$). The detailed post hoc test results related to $H1^*-H2^*$ values are available in Supplemental Materials. Based on the post hoc tests, all vowels except for /i/ showed higher $H1^*-H2^*$ in IDS than in ADS, implying that most vowels were typically breathier in IDS. The infant's age did not influence the values of the $H1^*-H2^*$ measure, and the post hoc test revealed that neither within IDS nor ADS did the age (4, 8, and 18 months) affect voice quality. Furthermore, age had no notable impact on the voice quality of any specific vowel.

The analysis of the other spectral tilt measure, $H1^*-A1^*$, revealed a similar relationship between voice quality and register as the $H1^*-H2^*$ measure (Table 3). The higher values of $H1^*-A1^*$ in IDS (compared to ADS) indicate a breathier voice. The detailed results of the post hoc tests related to the $H1^*-A1^*$ measure are reported in the Supplemental Materials. The test confirmed that this voice quality measure was systematically higher for most vowels in IDS than ADS, implying breathier voice production. Just as for $H1^*-H2^*$, the $H1^*-A1^*$ values for the vowel /i/ were an exception and showed an opposite trend. The infant's age did not affect either voice quality in general or the difference between the two registers. Moreover, the pairwise comparison within the realizations of each vowel also showed that age did not influence any of the vowels in this respect. To sum up, IDS was characterized by a breathier voice than ADS regardless of the infant's age.

We investigated the voice quality features of IDS not only with the spectral tilt measures ($H1^*-H2^*$, $H1^*-A1^*$) but also using CPP, which was introduced to quantify the periodicity of the signal. As expected, this measure also showed a difference between the two registers. The CPP values in IDS were generally lower than in ADS (Table 4), indicating that a less periodic voice was typical in this register, which is also a

Table 1. Mean and standard deviation (in brackets) of three voice quality measures and fundamental frequency in IDS ($n = 3125$) and ADS ($n = 3232$)

Register	$H1^*-H2^*$ M (SD)	$H1^*-A1^*$ M (SD)	CPP M (SD)	F0 (Hz) M (SD)
ADS	4.48 (5.23)	20.55 (7.16)	23.50 (4.23)	209.79 (50.84)
IDS	5.55 (5.64)	21.66 (7.60)	23.18 (4.08)	224.22 (56.64)

Table 2. The final linear mixed-effects regression model predicting H1*–H2* Regression model: H1*–H2* ~ Register + Age + Vowel + Register:Age Register:Vowel + Age:Vowel + (1 + Register+Age|Speaker), marginal $R^2 = 0.09$, conditional $R^2 = 0.18$

Term	b	SE	df	t	p
(Intercept)	2.72	0.48	44.43	5.7	< .001***
Register: IDS (reference level: “ADS”)	1.99	0.37	212.72	5.32	< .001***
Age: 8 months (reference level: “4 months”)	0.63	0.47	94.83	1.34	0.184
Age: 18 months (reference level: “4 months”)	−0.29	0.53	62.28	−0.54	0.589
Vowel: /ε/ (reference level: /a:/)	2.02	0.36	6280.04	5.59	< .001***
Vowel: /i/ (reference level: /a:/)	6.48	0.56	6282.45	11.55	< .001***
Vowel: /o/ (reference level: /a:/)	3.84	0.44	6284.19	8.8	< .001***
Vowel: /ɒ/ (reference level: /a:/)	1.25	0.37	6282.9	3.38	.001***
Register: IDS × Age: 8 months	−0.71	0.3	6290.14	−2.35	.019*
Register: IDS × Age: 18 months	−0.44	0.3	6291.23	−1.45	0.146
Register: IDS × Vowel: /ε/	−0.46	0.36	6280.23	−1.27	0.203
Register: IDS × Vowel: /i/	−2.83	0.56	6283.46	−5.06	< .001***
Register: IDS × Vowel: /o/	−0.3	0.43	6282.54	−0.68	0.494
Register: IDS × Vowel: /ɒ/	−0.59	0.37	6281.22	−1.61	0.108
Age: 8 months × Vowel: /ε/	−1.01	0.44	6280.1	−2.29	.022*
Age: 18 months × Vowel: /ε/	−0.41	0.44	6279.82	−0.93	0.351
Age: 8 months × Vowel: /i/	0.44	0.69	6283.2	0.64	0.523
Age: 18 months × Vowel: /i/	1.19	0.68	6280.52	1.74	0.082
Age: 8 months × Vowel: /o/	−1.42	0.53	6282.45	−2.69	.007**
Age: 18 months × Vowel: /o/	−0.76	0.54	6283.57	−1.42	0.155
Age: 8 months × Vowel: /ɒ/	−1.02	0.45	6280.99	−2.27	.023*
Age: 18 months × Vowel: /ɒ/	−0.2	0.45	6280.87	−0.44	0.656

Note: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

characteristic feature of breathy voice. The register, however, was found to interact with the vowel quality, but – as the pairwise tests revealed – this effect was clearly detectable only for the vowels /a:/ and /ε/. There, the CPP values were significantly lower in IDS than in ADS. The CPP values for the other vowels did not exhibit any significant register-specific differences (see the post hoc test results in the Supplementary Material). The infant’s age did not affect voice quality either in itself or in interaction with other factors; its inclusion as a parameter did not increase the model’s explanatory power, therefore, it was not incorporated into our final model. Generally, it can be concluded that, in terms of CPP, too, IDS showed the characteristics of a breathy voice, although this was not detectable in all vowels.

We also addressed the question of whether the fundamental frequency of speech differs in the two registers, similarly to studies investigating other languages (Benders, 2013; Narayan & McDermott, 2016). The results showed that the fundamental frequency

Table 3. The final linear mixed-effects regression model predicting H1*–A1* Regression model: H1*–A1* ~ Register + Age + Vowel + Register:Vowel + Age:Vowel + (1 + Vowel|Age|Speaker), marginal $R^2 = 0.02$, conditional $R^2 = 0.17$

Term	b	SE	df	t	p
(Intercept)	20.94	0.91	25.96	22.95	< .001***
Register: IDS (reference level: “ADS”)	2.31	0.39	6220.16	5.91	< .001***
Age: 8 months (reference level: “4 months”)	0.75	0.66	62.46	1.14	0.26
Age: 18 months (reference level: “4 months”)	−0.84	0.76	43.59	−1.11	0.275
Vowel: /ɛ/ (reference level: /a:/)	0.4	0.75	44.55	0.54	0.594
Vowel: /i/ (reference level: /a:/)	−1.07	1.08	51.45	−0.99	0.328
Vowel: /o/ (reference level: /a:/)	−0.06	1.04	35.37	−0.06	0.953
Vowel: /ɒ/ (reference level: /a:/)	0.32	0.86	36.5	0.37	0.715
Register: IDS × Vowel: /ɛ/	−1.48	0.49	6222.65	−3.01	.003**
Register: IDS × Vowel: /i/	−4.31	0.77	6236.90	−5.61	< .001***
Register: IDS × Vowel: /o/	−1.56	0.59	6229.64	−2.63	.009**
Register: IDS × Vowel: /ɒ/	−0.65	0.5	6220.84	−1.3	0.193
Age: 8 months × Vowel: /ɛ/	−1.54	0.6	6223.52	−2.54	.011*
Age: 18 months × Vowel: /ɛ/	−0.39	0.61	6225.81	−0.64	0.522
Age: 8 months × Vowel: /i/	0.36	0.95	6240.01	0.38	0.706
Age: 18 months × Vowel: /i/	2.11	0.94	6235.64	2.26	.024*
Age: 8 months × Vowel: /o/	−2.09	0.72	6236.48	−2.89	.004**
Age: 18 months × Vowel: /o/	−0.29	0.73	6240.08	−0.39	0.694
Age: 8 months × Vowel: /ɒ/	−1.8	0.62	6224.17	−2.91	.004**
Age: 18 months × Vowel: /ɒ/	−0.55	0.62	6223.44	−0.89	0.372

Note: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

of the vowels was higher in IDS than in ADS (Table 5). However, the children’s age did not affect the fundamental frequency of vowels.

4. Discussion

IDS can be distinguished from ADS by its numerous known semantic and acoustic features (Genovese et al., 2020; Wang et al., 2018). In agreement with the literature (Hilton et al., 2022; Kuhl et al., 1997), our analysis has also confirmed that the fundamental frequencies in IDS tend to be higher than those in ADS. Recent studies have drawn attention to another defining and potentially ubiquitous feature of IDS, breathy voice quality (Cheng et al., 2024; Chládková et al., 2019; Miyazawa et al., 2017). In our analyses, we investigated this relatively understudied property of IDS in Hungarian and evaluated whether and to what extent this feature changes as the infant ages.

Our results demonstrated that IDS typically exhibits a breathier voice than ADS, aligning with previous results on Canadian English-, Czech-, and Japanese-speaking

Table 4. The final linear mixed-effects regression model predicting CPP Regression model: CPP ~ Register + Vowel + Vowel:Register + (1 + Age|Speaker), marginal $R^2 = 0.03$ conditional $R^2 = 0.22$

Term	b	SE	df	t	p
(Intercept)	24.18	0.4	26.09	60.88	< .001***
Register: IDS (reference level: "ADS")	-1.02	0.21	6291.33	-4.82	< .001***
Vowel: /E/ (reference level: /a:/)	-0.03	0.19	6291.06	-0.14	0.889
Vowel: /i/ (reference level: /a:/)	-2.59	0.29	6291.70	-8.84	< .001***
Vowel: /o/ (reference level: /a:/)	-1.81	0.23	6292.07	-8.02	< .001***
Vowel: /6/ (reference level: /a:/)	-0.7	0.19	290.77	-3.67	< .001***
Register: IDS × Vowel: /ε/	0.63	0.27	6291.63	2.37	.018*
Register: IDS × Vowel: /i/	0.89	0.42	6291.40	2.14	.032*
Register: IDS × Vowel: /o/	1.37	0.32	6291.04	4.26	< .001***
Register: IDS × Vowel: /b/	0.84	0.27	6290.85	3.09	.002**

Note: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

Table 5. The final linear mixed-effects regression model predicting F0 Regression model: F0 ~ Register + Age + Vowel + (1 + Age|Speaker), marginal $R^2 = 0.04$, conditional $R^2 = 0.13$

Term	b	SE	df	t	p
(Intercept)	201.00	4.27	25.85	47.03	< .001***
Register: IDS (reference level: "ADS")	14.40	1.27	6304.05	11.31	< .001***
Age: 8 months (reference level: "4 months")	0.47	3.31	20.11	0.14	0.889
Age: 18 months (reference level: "4 months")	2.23	3.36	19.75	0.67	0.514
Vowel: /ε/ (reference level: /a:/)	17.03	1.85	6298.30	9.19	< .001***
Vowel: /i/ (reference level: /a:/)	21.27	2.88	6299.33	7.39	< .001***
Vowel: /o/ (reference level: /a:/)	10.19	2.23	6301.50	4.57	< .001***
Vowel: /b/ (reference level: /a:/)	-0.23	1.89	6297.75	-0.12	0.901

Note: ***: $p < .001$, **: $p < .01$, *: $p < .05$.

mothers (Cheng et al., 2024; Chládková et al., 2019; Miyazawa et al., 2017). To summarize, both spectral voice quality measures and those quantifying the periodicity of the phonation imply that Hungarian-speaking mothers also differentiate generally between the two registers in their voice quality. The cross-linguistic emergence of breathy voice in IDS may be motivated by several factors. First, when communicating with each other, people tend to adapt the characteristics of their speech to the other person (Bernhold & Giles, 2020). Such speech and phonetic adaptation may be driven not only by the perceived acoustic cues in the situation but also by previous experiences or stereotypes. Indeed, some of the features of IDS (e.g., higher fundamental frequency, lower speech rate) are also typical traits of children's speech (Payne et al., 2012; Zhang, 2021). Clearly, the presence of these features of IDS is motivated by numerous other factors as well, yet, in the case of voice quality, it is important to note that the children's phonation is typically

breathier than that of adults (Kent et al., 2021; Zhang, 2021). The acoustic differences in the fundamental frequency and voice quality can be derived from anatomical differences between adults and children, namely in vocal fold length and thickness (Zhang, 2021). These body size-dependent characteristics may also carry additional culture-specific meaning. In British English, women with breathier voices were perceived as happier and more attractive, and men with breathier voices were rated as friendlier and happier. (Noble & Xu, 2011; Xu et al., 2013). Certainly, the expression of emotions and the representation of personality traits through voice quality can be largely culture-dependent. In interpreting the present results, it is important to emphasize that Hungarian native speakers also tend to associate breathy voice with positive emotions (Bartók, 2018). While the use of breathy voice in IDS may serve to express positive emotions or reflect phonetic adaptation, it seems less likely that speakers would use breathy voice for a direct pedagogical purpose. However, the speaker's change of register from ADS to IDS may also serve as a signal to children, indicating that relevant information will follow and is worthy of their attention. Thus, it cannot be ruled out that this attention-raising aspect of IDS may facilitate language acquisition (Senju & Csibra, 2008). Studies have shown that a happy sound produced with a breathy voice maintained an infant's attention more than a happy sound with a modal voice or a neutral sound with any voice quality (Kao et al., 2022). This attention-maintaining function of breathy voice may, therefore, indirectly contribute to language development.

The acoustic features of IDS may be modified as the child ages; thus, we also investigated whether the breathiness of the voice changes with the infant's age. Generally, our results indicated that child's age – up to 18 months – had no effect on the mother's voice quality. This is in agreement with the earlier results that even at 2 years, the voice quality difference between the two registers was clearly detectable (Miyazawa et al., 2017). Some acoustic features of IDS, such as speech rate and fundamental frequency, change in the first few years of a child's life, while others, like vowel space area and F_0 variability, remain relatively stable (Cox et al., 2023). It appears that the breathy voice quality of Hungarian IDS belongs to the latter category, showing no detectable changes over the first 18 months. This implies that parents are, in a certain sense, motivated to maintain this feature of IDS. The need for continuous positive emotion expression or the constant phonetic adaptation can be the reason why the breathy voice remains a stable, permanent element of the IDS repertoire, regardless of the child's age.

5. Limitations and future directions

The research reported in the present study had several limitations, which may be worth extending and further investigating in the future. First, the results of our measurements indicated that the phonation of IDS is indeed characterized by a breathier voice, while the analyses also revealed differences between the various vowels. This is not surprising, given the fact that vowel quality largely influences the values of the studied voice quality measures, even though the spectral tilt measures are corrected for vowel formants (Garellek & Esposito, 2023). In our material, the vowel /i/ was the one that occurred in the smallest number of phonetic contexts, so the investigation of further factors would not be feasible in more complex linear mixed models than those applied here. Previous literature on IDS has reported voice quality differences between vowels (Cheng et al., 2024), but a detailed analysis of which vowels differ, in what contexts, and to what extent in the two registers is left for future targeted research. It should be noted that not all

syllables in speech need to be breathy to create a certain general perception in the listener. Secondly, the voice quality measure related to the periodicity also did not exhibit a difference for all vowels between the two registers. Note that based on this particular measure (CPP) alone, creaky and breathy voices cannot be distinguished, as both of these voice quality types yield lower CPP values than modal voice. In Hungarian, it is known that certain prosodic and phonetic positions may trigger the emergence of creaky voice (Markó, 2013; thus, these situations were excluded from our analyses. However, Hungarian has a rather large ratio of syllables pronounced with creaky voices for which we currently do not have an explanation (Grácz et al., 2017). This peculiarity of Hungarian may affect the interpretation of CPP-related results. Further research is needed to clarify how the rarely observed creaky voice in ADS is realized in IDS. Finally, this study investigated read sentences integrated into spontaneous storytelling. Future research should examine the voice quality of spontaneous IDS and compare the acoustic characteristics across different speech styles.

6. Conclusion

To conclude, this study revealed that Hungarian mothers' speech addressed to their infants exhibits a breathier voice. These findings are in line with previous studies of Japanese, Canadian English, and Czech IDS (Cheng et al., 2024; Chládková et al., 2019; Miyazawa et al., 2017). The breathy voice was found to be continuously present in the mothers' IDS throughout the first 18 months of the infants' age. The appearance of a breathy voice may be motivated by the mother's phonetic adaptation to the child's speech characteristics or by expressing positive emotions towards the child. Among the acoustic characteristics of IDS, voice quality is a relatively underexplored feature that can help increase and maintain the child's attention. Whether and to what extent this function can also contribute to the child's language development remains an open question for future research.

Abbreviations

ADS adult-directed speech
IDS infant-directed speech

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S0305000926100506>.

Acknowledgements. We would like to thank all the families who have contributed to this study. We gratefully acknowledge the work of Katalin Pirsell and Luca Garai, who contributed to the annotation of the sentences.

Funding statement. This study was supported by the Hungarian National Research, Development and Innovation Office (grants PD134775 and K115385), and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (BO/160/25).

Competing interests. The authors declare no competing interests.

References

- Anikin, A. (2020). A moan of pleasure should be breathy: The effect of voice quality on the meaning of human nonverbal vocalizations. *Phonetica*, 77(5), 327–349.
- Bartók, M. (2018). A gégeműködés variabilitása az érzelmkifejezés függvényében. *Beszédkutatás*, 26(1), 30–62.

- Bartoń, K.** (2024). Mumin: Multi-model inference [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=MuMIn> (R package version 1.48.4)
- 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.
- Bergelson, E., & Swingle, D.** (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, **109**(9), 3253–3258.
- Bergelson, E., & Swingle, D.** (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, **11**(4), 369–380.
- Bernhold, Q. S., & Giles, H.** (2020). Vocal accommodation and mimicry. *Journal of Nonverbal Behavior*, **44**(1), 41–62.
- Boersma, P., & Weenink, D.** (2019). Praat: doing phonetics by computer. <http://www.praat.org/>.
- Cheng, A., McClay, E., & Yeung, H. H.** (2024). An exploration of voice quality in mothers speaking Canadian english to infants. *Language Learning and Development*, **20**(4), 279–296.
- Chládková, K., Černá, M., Paillereau, N., Skarnitzl, R., & Oceláková, Z.** (2019). Prenatal infant-directed speech: Vowels and voice quality. In *Proceedings of ICPHS* (pp. 1525–1529). ICPHS
- 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*, **94**(6), 1672–1696.
- Garellek, M.** (2019). The phonetics of voice 1. In *The Routledge Handbook of Phonetics* (pp. 75–106). Routledge.
- Garellek, M.** (2022). Theoretical achievements of phonetics in the 21st century: Phonetics of voice quality. *Journal of Phonetics*, **94**, 101155.
- Garellek, M., & Esposito, C. M.** (2023). Phonetics of white hmong vowel and tonal contrasts. *Journal of the International Phonetic Association*, **53**(1), 213–232.
- Genovese, G., Spinelli, M., Lauro, L. J. R., Aureli, T., Castelletti, G., & Fasolo, M.** (2020). Infant-directed speech as a simplified but not simple register: A longitudinal study of lexical and syntactic features. *Journal of Child Language*, **47**(1), 22–44.
- Gráci, T. E., Markó, A., & Takács, K.** (2017). Az irreguláris zöngje megjelenése mondatfelolvasásban: Tizenéves és felnőtt beszélők adatainak összehasonlítása. *Alkalmazott Nyelvtudomány*, **17**(1), 1–15.
- Háden, G. P., Mády, K., Török, M., & Winkler, I.** (2020). Newborn infants differently process adult directed and infant directed speech. *International Journal of Psychophysiology*, **147**, 107–112.
- Harmati-Pap, V., Vadász, N., Tóth, L., & Kas, B.** (2024). Patterns of lexical and syntactic adjustment in early infant-directed speech related to language development in Hungarian. *Clinical Linguistics & Phonetics* **39**, 1–29.
- Hillenbrand, J., Cleveland, R. A., & Erickson, R. L.** (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech, Language, and Hearing Research*, **37**(4), 769–778.
- Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., ... Mehr, S. A.** (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour*, **6**(11), 1545–1556.
- Iseli, M., Shue, Y.-L., & Alwan, A.** (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America*, **121**(4), 2283–2295.
- 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.
- Kalashnikova, M., Carignan, C., & Burnham, D.** (2017). The origins of babytalk: Smiling, teaching or social convergence? *Royal Society Open Science*, **4**(8), 170306.
- Kao, C., Sera, M. D., & Zhang, Y.** (2022). Emotional speech processing in 3-to 12-month-old infants: Influences of emotion categories and acoustic parameters. *Journal of Speech, Language, and Hearing Research*, **65**(2), 487–500.
- Kawahara, H., De Cheveigne, A., Banno, H., Takahashi, T., & Irino, T.** (2005). Nearly defect-free f0 trajectory extraction for expressive speech modifications based on straight. In *Interspeech* (pp. 537–540). International Speech Communication Association (ISCA).
- Kawahara, S., & Shinya, T.** (2008). The intonation of gapping and coordination in Japanese: Evidence for intonational phrase and utterance. *Phonetica*, **65**(1–2), 62–105.
- Kent, R. D., Eichhorn, J. T., & Vorperian, H. K.** (2021). Acoustic parameters of voice in typically developing children ages 4–19 years. *International Journal of Pediatric Otorhinolaryngology*, **142**, 110614.

- Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347.
- Kohári, A., Reichel, U. D., Szalontai, Á., & Mády, K. (2024). A magánhangzók zöngeminősége frázisok határán a dajkanyelvi beszédben. *Alkalmazott Nyelvtudomány*, 24(1), 33–52.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., ... 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(1), 1–26.
- Ladefoged, P. (1973). The features of the larynx. *Journal of Phonetics*, 1(1), 73–83.
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). *Emmeans: Estimated marginal means, aka least-squares means*. [Computer software manual]. Retrieved from <https://cran.r-project.org/web/packages/emmeans/index.html> (R package version 1.3.4).
- Mády, K., Gyuris, B., Gärtner, H.-M., Kohári, A., Szalontai, A., & Reichel, U. D. (2022). Perceived emotions in infant-directed narrative across time and speech acts. In Frota, S., Cruz, M., & Vigário, M. (Eds.), *Proceedings 11th international conference on speech prosody* (pp. 590–594). International Speech Communication Association (ISCA). <https://doi.org/10.21437/SpeechProsody.2022-120>.
- 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.
- Markó, A. (2013). Az irreguláris zöngé funkciói a magyar beszédben. ELTE Eötvös Kiadó.
- Markó, A., Deme, A., Bartók, M., Grácz, T. E., & Csapó, T. G. (2018). Speech rate and vowel quality effects on vowel-related word-initial irregular phonation in Hungarian. In *Challenges in analysis and processing of spontaneous speech*. MTA Nyelvtudományi Intézet.
- McClay, E. K., Cebiglu, S., Broesch, T., & Yeung, H. H. (2022). Rethinking the phonetics of baby-talk: Differences across Canada and Vanuatu in the articulation of mothers' speech to infants. *Developmental Science*, 25(2), e13180.
- 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.
- 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.
- Noble, L., & Xu, Y. (2011). Friendly speech and happy speech-are they the same? In *ICPHS* (pp. 1502–1505). University College London.
- Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. d. M. (2012). Measuring child rhythm. *Language and Speech*, 55(2), 203–229.
- R Core Team. (2021). *R: A language and environment for statistical computing* [computer software manual]. Vienna, Austria: R Core Team.
- 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.
- Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., ... Cohen, D. (2013). Motherese in interaction: At the cross-road of emotion and cognition? (a systematic review). *PLoS One*, 8(10), e78103.
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18(9), 668–671.
- Shue, Y.-L., Keating, P., Vicens, C., & Yu, K. (2011). Voicesauce: A program for voice analysis. In *Proceedings of the Seventeenth International Congress of Phonetic Sciences* (pp. 1846–1849). International Phonetic Association.
- 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.
- Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, 17(4), 432–444.
- Wang, Y., Houston, D. M., & Seidl, A. (2018). Acoustic properties of infant-directed speech. *The Oxford Handbook of Voice Perception*, 1, 93–116.

- Winter, B.** (2019). *Statistics for Linguists: An Introduction Using R*. Routledge. <https://doi.org/10.4324/9781315165547>
- Xu, Y., Lee, A., Wu, W.-L., Liu, X., & Birkholz, P.** (2013). Human vocal attractiveness as signaled by body size projection. *PLoS One*, **8**(4), e62397.
- Zhang, Z.** (2021). Contribution of laryngeal size to differences between male and female voice production. *The Journal of the Acoustical Society of America*, **150**(6), 4511–4521.