Journal of Nonverbal Behavior (2024) 48:513–537
https://doi.org/10.1007/s10919-024-00462-z

ORIGINAL PAPER

Vocal Emotional Expressions in Mothers with and without a History of Major Depressive Disorder

Emma Ilyaz[1] · Xin Feng[2] · Xiaoxue Fu[3,4] · Eric E. Nelson[4,5] · Michele Morningstar[1,4,6]

Accepted: 11 April 2024 / Published online: 25 April 2024

Abstract
Depression is associated with alterations in prosody when speaking (e.g., less variation in pitch, slowed speech rate), but less is known about its impact on *emotional* prosody. This is particularly important to investigate in parent–child contexts, as parental expression of emotion may contribute to the intergenerational transmission of depression risk. The current study asked mothers of preschool-aged children (with and without a history of major depressive disorder during their child's lifetime) to produce child-relevant sentences in neutral, angry, and happy tones of voice. We examined whether groups' portrayals were acoustically or perceptually different, in speech analyses and listener ratings. Mothers with a history of depression expressed happiness with less range in pitch and a slower speech rate (slower, more monotonous voice) than mothers with no history of depression. Across groups, happy exemplars with less range in pitch were rated as less emotionally intense, recognizable, and authentically happy by listeners; slower speech rate was associated with opposite perceptual ratings. However, listeners' ratings did not differ by depression group as a whole. Results suggest that a history of depression may influence maternal vocal expression of happiness, but that its impact on listeners' perceptions may depend on mothers' idiosyncratic use of acoustic cues.

Keywords Affective prosody · Emotional expression · Emotion perception · Depression · Maternal mental health

Introduction

Depression has been associated with atypical nonverbal communication—that is, differences in the facial expressions, postures and body movements, and vocal prosody used during interpersonal interactions. For instance, compared to control groups without depression, individuals with depression show fewer facial expressions of happiness (Girard et al., 2013, 2014), make fewer facial movements overall (Gehricke & Shapiro, 2000), and walk more slowly and with a slumped posture (Feldman et al., 2020; Michalak et al., 2009). Nonverbal communication differences have also been identified in prosodic patterns of speech. Acoustic analyses have shown that individuals with depression are more monotone when they speak, using less emphasis and inflections (Alpert et al., 2001) and an overall reduced variation in their pitch (Darby et al., 1984). In addition, depression scores are negatively

Extended author information available on the last page of the article correlated with speaking rate (Cannizzaro et al., 2004) and overall intensity (loudness) of speech (Darby et al., 1984). As with facial expressions (Cohn et al., 2009), differences in patterns of vocal expression are a reliable indicator of the presence of depressive symptoms (e.g., McGinnis et al., 2019; Pan et al., 2019; Silva et al., 2021). Conversational partners can also detect depression severity based on vocal cues, and adjust their own speech accordingly (e.g., speaking with greater range in pitch to more severely depressed individuals; Yang et al., 2012). As such, the alterations in vocal prosody associated with depression might impact the interpersonal dynamics of social interactions.

The effects of depression on vocal expression patterns might be particularly important to consider within the parent–child relationship. Maternal depression is a well-known risk factor for depression in offspring (Goodman, 2020). Parental histories of depression can have long-term developmental impacts: for instance, a history of depression in mothers was associated with internalizing symptoms in their adolescents, even when controlling for concurrent symptoms in the mother (O'Connor et al., 2017). Atypical expressions of emotion have been hypothesized to be a mechanism through which intergenerational risk for depression is transmitted. For instance, maternal depression has been linked to greater negative emotionality (e.g., expression of anger or frustration, criticism of the child) in mother–child interactions, which was associated with greater internalizing and externalizing symptoms (Gravener et al., 2012) and poorer emotion recognition (Kujawa et al., 2014) in young offspring of depressed mothers. In addition, low levels of maternal encouragement and a tendency to 'dampen' their children's positive emotion (i.e., mothers' self-report of scolding, punishing, disapproving of, or feeling irritated by their children's positive emotion) was associated with blunted ventral striatum response to reward—a neurobiological index of risk for depression in youth (Kujawa & Burkhouse, 2017)—in children at high familial risk for depression (Morgan et al., 2022). Because children's understanding of emotion is at least partly socialized by their parents (e.g., Barrett & Campos, 1987; Morris et al., 2007; Capatides & Bloom, 1993), heightened negativity and blunted positivity in parents' interactions with their children may influence how children learn to regulate their emotions (Morris et al., 2007, 2017; Thompson, 1991), which may contribute to their risk for depression (Loechner et al., 2020; Silk et al., 2006).

While previous research has focused on the impact of maternal depression on mothers' verbal expressions of emotion or behaviour in interactions with their child (e.g., Azak & Raeder, 2013; Feldman et al., 2009; Gravener et al., 2012; Hoffman et al., 2006; Hooper et al., 2015), the experience of depression may also disrupt parents' *nonverbal* expression of emotion (e.g., prosodic patterns in speech). Maternal prosody is an important component of parent–child interactions: for instance, a study by Kolacz and colleagues (Kolacz et al., 2022) found that the acoustic characteristics of mothers' voices were related to changes in their infants' emotional regulation during a still-face paradigm. Parents with depression have also been found to use less prosodic variations in their infant- or child-directed speech (Kaplan et al., 2002; Lam-Cassettari & Kohlhoff, 2020), a form of exaggerated sing-song speech that is thought to promote children's learning and social interaction with their parent (Saint-Georges et al., 2013). Given this, it is crucial to investigate not only how mothers' experience of depression impacts what they say when expressing emotion—but also *how* they say it.

Current Study's Goals and Hypotheses

The current study examined whether depression was associated with variations in the vocal characteristics of maternal emotional prosody, in mothers with and without a history of major depression during their child's lifetime. Mothers were asked to perform various child-relevant sentences in a neutral, angry, and happy tone of voice. Most studies examining nonverbal communication in individuals with depression have described characteristics of free, unconstrained speech in naturalistic or semi-natural environments. In contrast, less is known about how depression may impact the *purposeful* nonverbal expression of emotional states that are generated on demand. Speakers can purposefully modulate their voice to convey specific emotions, in ways that inform others' perception of their emotional intent (Bachorowski, 1999); these emotional prosody patterns deviate meaningfully from speakers' natural (non-emotional) way of speaking (Johnstone & Scherer, 2000; Juslin & Laukka, 2003). Understanding how the experience of depression may impact the purposeful manipulation of these vocal cues is important for several reasons. First, recording vocal expressions in a controlled experimental environment allows for quantitative analysis of prosody that controls for natural variations in lexical content, context, and environment. Second, generating a simulated emotional expression requires the explicit manipulation of prosodic cues and demands a certain level of emotion regulation, which may be disrupted by depression (Joormann & Stanton, 2016). Finally, the upregulating of prosodic cues associated with emotionality (i.e., exaggerating vocal cues associated with the expression of emotion) is a common occurrence in parent–child dyadic interactions, to engage joint attention and social communication (e.g., in infant-directed speech; Spinelli et al., 2017; Trainor et al., 2000). As such, understanding parental use of *emotional prosody* complements existing work on naturalistic speech patterns to better characterize child-relevant emotional communication patterns.

We investigated a) whether there were differences in how mothers with a recent history of depression (MHD) and mothers with no history of depression (MND) conveyed emotions vocally in performed speech (using speech analysis to quantify acoustic characteristics of their voice), and b) whether these differences were associated with listeners' perception of their emotional speech (based on subjective ratings by independent judges). A large body of research has determined that happiness is typically communicated with high mean pitch and high variation in pitch, as well as high acoustic intensity and fast speech rate (i.e., higher-sounding, "up and down" speech, spoken quickly and relatively loudly); anger is typically expressed with high intensity and fast speech rate (i.e., speaking loudly and quickly; see reviews by Johnstone & Scherer, 2000; Juslin & Laukka, 2003). Given previous findings of altered vocal cues in the naturalistic speech of individuals with depression, we hypothesized that MHD would communicate anger and happiness in atypical ways. Because depression has been characterized by blunted affect (Bean et al., 2022; Werner-Seidler et al., 2013), we expected that mothers in the MHD group would express both emotions with lower mean pitch and smaller pitch range (for happiness), and lower acoustic intensity and a slower speech rate (for both) compared to MND group. We also hypothesized that vocal expressions of happiness and anger in MHD would be rated as being less emotionally intense, less recognizable, and less authentic than those of MND by independent judges (who were blind to speakers' depression group status). However, given that current conceptualizations of depression also describe increases in negative affect (Clark & Watson, 1991; Forbes et al., 2004; Joormann & Vanderlind, 2014), we also considered that mothers in the MHD group may instead express anger with higher acoustic intensity and a faster speech rate compared to the MND group—and that these portrayals may be deemed better exemplars of anger by independent judges. Delineating whether and how depression impacts the purposeful expression of emotion—and how it is perceived by others—helps further our understanding of socio-communicative processes in depression and may offer a target for clinical intervention.

Method

Participants

Eighty-one mothers and their 3- to 4-year-old children were recruited from the community in a large Midwestern American city ($M_{\text{age mother}}=35.53$ years, $SD=4.66$, age range = 25–47 years; $M_{\text{age child}}=4.04$ years, $SD=0.16$) as part of a larger study on maternal history of depression and children's autobiographical memory. Mothers with and without a history of depression during their child's lifetime (21 years of age or older) were recruited from the local university, health clinics, school districts, daycare centers/preschools, community centers, and through online advertisements. Mothers were screened using the Structural Clinical Interview for DSM-5 Disorders (First et al., 2016) to determine depression status (31% of interviews were coded by a second rater; kappa = 0.92 for reliability of major depression disorder diagnosis). Exclusion criteria for mothers included lifetime psychosis or bipolar disorder, or substance use disorder within the past 6 months. Mothers were considered part of the MHD group if they had a major depressive episode during their child's lifetime ($n=39$); mothers without past or current diagnoses of depression were included in the MND group ($n=42$). The two groups differed in concurrent symptom load: MHD self-reported significantly higher depression symptoms ($M=18.64$, $SD=11.09$) on the Beck Depression Inventory (Beck et al., 1996) than MND ($M=5.14$, $SD=5.13$), $t(79)=7.11$, $p<0.001$. Mothers in the MHD group ranged from having 1–10 + episodes of major depression during their child's lifetime, with 10 mothers only having 1 major depressive episode and 6 mothers having 10 + episodes of major depression during their child's lifetime. At the time of the study, 10 mothers in the MHD depression group were in a current major depressive episode.

According to self-reports, 83% of mothers were white, 12% were Black, and 5% were of mixed racial backgrounds; 5% reported Hispanic or Latino ethnicity. Mean income was $95,000 (ranging from $5,000-$175,000). The majority (89%) of the sample reported having a college degree; 68% were currently employed; 90% cohabited with a spouse or partner. MHD reported slightly less income than MND, $t(79)=2.74$, $p=0.01$, but there were no group differences in mothers' age, $t(79)=0.51$, $p=0.61$, children's age, $t(79)=-0.17$, $p=0.86$, mothers' race, $\chi^2(2)=4.10$, $p=0.13$, education, $\chi^2(5)=5.70$, $p=0.34$, employment status, $\chi^2(3)=2.24$, $p=0.53$, or marital status, $\chi^2(3)=6.15$, $p=0.11$. All mothers were fluent in written and spoken English.

The study was approved by institutional ethics review boards. Mothers gave informed consent for their participation, and received a $75 gift card for their time. Mothers were informed of the study procedures, but not of study hypotheses.

Procedure

Collection of Recordings

During a lab visit, mothers completed a 20-minute voice recording task during which they were first asked to say twenty sentences (Appendix) in a neutral tone (e.g., "Put your shoes away"). They were then asked to portray the same sentences in an angry tone of voice, followed by a happy tone of voice. (There was a short break between emotion types to minimize carry-over effects across conditions.) This 'standard content paradigm' is a well-established method for examining acoustic characteristics associated with emotional prosody, while controlling for speech content (see reviews by Juslin & Laukka, 2003; Scherer, 2003). Research assistants were blind to the mother's depression group assignment. Neither coaching nor feedback was provided to participants during the task. Mothers were only instructed to portray emotions prototypically and with high emotional intensity (e.g., "purposefully happy sounding", "purposefully angry sounding"), without specifying an intended audience. Sentences were designed to be generally neutral in content, relevant to parent–child interactions, and feasibly expressed in either a positive (happy) or negative (angry) tone. Mothers' voices were recorded using a Zoom H2N recorder (44.1 kHz sampling rate) placed 5 cm away from them at a 45º angle from the speaker's mouth (controlling for variations in acoustic intensity alongside normalization procedures described below; Pell et al., 2009). This procedure yielded a total of 4860 recordings (81 mothers $\times$ 20 sentences $\times$ 3 emotional tones of voice).

Listeners' Perception of Recordings

To characterize how these recordings were perceived by listeners, 452 listeners ($M_{age} = 21.95$ years, $SD = 7.09$, age range = 17–58 years; 79% identifying as women, 20% as men, < 1% as other genders, and < 1% preferring not to answer) were recruited from a participant pool at a mid-sized Canadian university. Listeners self-reported the following ethnic origins (selecting from national census categories): 55% reported European origins, 25% Asian origins, 3% African origins, 2% Caribbean origins, 2% Latin and Central American origins, 1% North American Aboriginal origins, 5% other North American origins, 4% another ethnic origin, and 3% did not report their ethnic origins. All listeners reported normal hearing and fluency in English; 94% of participants reported that English was the language that they were more comfortable speaking, with another < 1% of participants reporting French and 6% reporting another language as their dominant language. Listeners were given a 1% bonus credit towards their psychology courses for participating. Forty-two listeners from the recruited sample were excluded for failing a headphone check (see below); the final sample thus comprised 410 listeners. Listeners provided informed consent to participation as part of a larger study on the perception of emotional voices.

To feasibly obtain listener ratings for each of 4860 recordings, we split recordings obtained from mothers into separate sets of 60 recordings each. Each set included an equal number of happy, angry, and neutral recordings (20 each), with half of these being produced by mothers in each group. Sets were randomly assigned to listeners. As a result, each recording was rated by 4–9 listeners (average 5.95). Recordings were not low-pass filtered, and thus contained semantic information; however, each listener heard each sentence in all emotional tones (i.e., each sentence spoken in a happy, angry, and neutral tone). Statistical models (described below) controlled for sentence-specific effects, as well as random-effects of listeners.

The perception task was completed online through Pavlovia (Pierce et al., 2019). Listeners were required to wear headphones, verified using a headphone check task (developed by Woods et al., 2017). In this task, listeners were asked to identify the quietest sounding tone in a series of 3 (presented out of phase); although easy to do with functioning headphones, this task is nearly impossible without them (Woods et al., 2017). The headphone check thus allowed us to ensure that participants were wearing headphones, and that their headphones were functioning properly. Only participants who passed the headphone check (i.e., correctly identified the quietest sound in a minimum of 5/6 trials) completed the rest of the perception task.

To obtain a measure of listeners' emotion recognition accuracy for each recording, listeners were first presented with each recording in the set (in randomized order) and asked to select the speaker's intended emotion from 3 label choices (anger, happiness, or neutral). Second, to obtain listeners' ratings of the quality of each emotional expression, listeners were again presented with the angry and happy recordings (blocked by emotion type, order counterbalanced across listeners; recordings within each emotion type presented in a new randomized order) and informed of the intended emotion for each. They were then asked to rate the emotional intensity, recognizability, and authenticity of the recording's intended emotion, on 1–4 scales (1 = not at all, 4 = very much so; based on scales used in Banse & Scherer, 1996; Scherer & Ellgring, 2007; Morningstar et al., 2018). The second set of ratings, in which the intended emotion was communicated to listeners prior to their judgments of each recording, was obtained to ensure that errors in identifying the emotion type (i.e., as in the first part of the task) did not colour listeners' interpretation of how emotionally intense, recognizable, and authentic each recording was. Neutral recordings were not included in the second part of the task, given the inherent difficulty in rating the emotional properties of putatively non-emotional recordings.

Analyses

Speech Analysis

All voice recordings were edited using Audacity (Audacity, 2020) to correct for overall volume inconsistencies, noise reduction, and echo. To ensure quality control, research assistants listened to each recording with headphones to remove recordings of unsuitable quality (e.g., participant said wrong words, room noise interrupted participant's speech, etc.). As a result of this procedure, a total of 70 individual sentence recordings (1% of sentence recordings; 20 from MND, 50 from MHD) were removed from the sample before statistical analyses. Speech analysis was performed using Praat (Boersma & Weenick, 2001) to extract information about pitch (fundamental frequency[1]), acoustic intensity, and speech rate variables.

Acoustic Variables Three pitch measures (mean, minimum, and maximum pitch; measured in Hertz (Hz)) and three intensity measures (mean, minimum and maximum intensity; measured in decibels (dB)) were computed across the whole sentence, following procedures outlined in Morningstar et al. (2017).[2] Resulting values were assessed for potential pitch doubling or halving; all values were considered within

[1] Pitch is the perceptual correlate of fundamental frequency (F0); in the voice, this is the frequency at which vocal folds vibrate when speakers produce speech (Lee & Humes, 2012). We refer to measures of F0 as pitch for ease of interpretation.

[2] Pitch measures were obtained across the entirety of the waveform using autocorrelation, time step = 0.0 s, pitch floor = 75 Hz, pitch ceiling = 600 Hz (to apply to higher-pitched voices). Intensity measures were also obtained across the whole waveform, minimum pitch setting = 100 Hz, time step = 0.0 s.

typical ranges for women (e.g., Goy et al., 2013). Speech rate was calculated by dividing the number of syllables within the sentence by the sentence's duration in seconds (s). The resulting speech rate variable was measured in syllables/second (syllables/sec), where higher values reflect faster speech rate. Acoustic variables were measured across each individual recording; although admittedly coarse measures, this procedure enables comparison with decades of research on acoustic correlates of emotionality in the voice (Pell et al., 2009; Banse & Scherer, 1996; Johnstone & Scherer, 2000; Spackman et al., 2009).

Normalization to Baseline To account for individual differences in speaking patterns, all emotional voices were normalized to the speaker's 'resting voice', using best-practice methods (outlined in Pell et al., 2009). Each speaker's resting frequency was computed by averaging their minimum pitch across the 20 sentences they produced in the neutral condition. The same computation was done to obtain a resting value for intensity (i.e., an average of minimum intensity across the speaker's 20 sentences in the neutral condition).

Mean pitch for each speaker's sentence (60 sentences per speaker) was then normalized to that speaker's resting voice by subtracting the participant's computed resting frequency for pitch from the sentence's observed mean pitch and dividing it by the speaker's resting frequency. The following equation was used: normalized pitch mean = (observed mean pitch — resting frequency) / resting frequency. This process was also conducted for the minimum and maximum pitch of each sentence. Normalized pitch range was then computed by subtracting the normalized minimum pitch from the normalized maximum pitch. Normalized mean, minimum, maximum, and range of intensity were obtained using the same procedure. Normalized values represent the proportional distance from a speaker's natural way of speaking (Pell et al., 2009), thereby controlling for potential differences in speakers' vocal tracts and natural frequency levels (e.g., avoiding group differences due to natural differences in pitch or small variations in microphone placement). All analyses used normalized values for pitch and intensity variables.[3]

Statistical Analyses

Speech Variables Using the lmerTest package (Kuznetsova et al., 2017) in R (R Core Team, 2021), linear mixed-effect models were performed to investigate the effect of mothers' depression group (between-subjects factor: MHD vs. MND, with MND as the reference category) and emotion (within-subjects factor: angry, happy, and neutral, with neutral as the reference category) on each acoustic cue (mean pitch, pitch range, mean intensity, intensity range, and speech rate). Sentence type (i.e., which of the 20 standardized sentences was spoken in each recording) was entered as a control variable, to account for variations in linguistic content across sentences. Random per-speaker and

[3] To ensure that normalizing procedures did not obscure individual differences relevant to maternal history of depression, we recomputed analyses using raw (non-normalized values). All results were identical to those presented in text.

speaker-by-emotion intercepts were also included in the model.[4] A separate model was built for each acoustic cue. Models were built sequentially: we first added the random speaker-by-emotion intercept (which significantly improved model fit for each acoustic cue; all *p*s < 0.001) to the base model with only the fixed and random per-speaker intercept, followed by the fixed effects of emotion, sentence, depression group, and the interaction of depression group and emotion, in sequence. Estimated *p* values (reported in text) for all fixed effects were obtained from likelihood ratio tests of the full model with the effect in question against the previous model in this model-building sequence.

Listeners' Perception of Emotional Recordings Given the online nature of the perception task, we first screened for potential response sets among listeners. We computed how many times each listener used a given response for each rating scale (e.g., how many times someone responded '4' on a given scale). Listeners who used a given response more than 2 standard deviations more often than the sample mean (suggestive of a response pattern, e.g., indiscriminately answering '4' on the rating scale) were removed from the sample prior to analysis. This procedure resulted in the removal of 8 listeners (2% of the 410 listeners who passed the headphone check).

Listeners' recognition accuracy (0 or 1 for each recording, from the first part of the perception task) and perceptual ratings of emotional intensity, recognizability, and authenticity (continuous values from 1–4, from the second part of the perception task) were obtained for each angry and happy recording. We fit linear mixed-effect models to test whether emotional recordings produced by MHD were recognized or perceived differently than those produced by MND, and whether this effect varied by emotion type. Models included both per-speaker and per-listener random intercepts.[5] Separate models were fit for listeners' recognition accuracy, ratings of emotional intensity, ratings of recognizability, and ratings of authenticity. Models were built sequentially: we first added the fixed effect of depression group to the base model with the fixed and random intercepts, followed by the interaction of depression group and emotion type. Estimated *p* values (reported in text) for each fixed effect were obtained from likelihood ratio tests of the full model with the effect in question against the previous model in this model-building sequence. R scripts for all analyses can be found on OSF (https://osf.io/mv4jp/).

Results

Group Differences in Acoustic Cues Associated with Emotional Expression

Parameter estimates (e.g., B, *t*) for all models are presented in Table 1. There were significant main effects of emotion type on all acoustic variables, *p*s ≤ 0.001 (Table 1), across both depression groups. Post-hoc comparisons indicated that happiness was expressed with greater mean pitch and pitch range than anger, which was higher in these pitch variables than neutral (*p*s < 0.001). Anger was higher in intensity range than happiness, which was

[4] The R equation was as follows: lmer(AcousticCue ~ 1 + DepressionGroup*Emotion + Sentence + (1 | SpeakerID) + (1 | SpeakerID:Emotion).

[5] The R equation was as follows: lmer(Rating ~ 1 + DepressionGroup*Emotion + (1 | SpeakerID) + (1 | ListenerID).

Table 1 Parameter estimates for the effects of depression group and emotion type on acoustic cues

Acoustic cue	Parameter	Estimate	*SE*	*t*	*p*	η^2
Normalized pitch mean	Intercept	0.55	0.05	11.19	< .001***	
	Emotion: angry	0.42	0.04	10.02	< .001***	.80
	Emotion: happy	0.80	0.04	19.30	< .001***	
	Depression group: with depression	-0.09	0.07	-1.31	.19	.04
	Emotion: angry x Depression group: with depression	0.02	0.06	0.33	.75	.01
	Emotion: happy x Depression group: with depression	-0.06	0.06	-0.98	.33	
Normalized pitch range	Intercept	1.14	0.08	13.50	< .001***	
	Emotion: angry	0.34	0.07	5.16	< .001***	.75
	Emotion: happy	1.10	0.07	16.55	< .001***	
	Depression group: with depression	-0.20	0.11	-1.84	.07	.03
	Emotion: angry x Depression group: with depression	0.20	0.10	2.03	.04*	.04
	Emotion: happy x Depression group: with depression	-0.04	0.10	-0.45	.66	
Normalized intensity mean	Intercept	0.81	0.05	16.50	< .001***	
	Emotion: angry	0.02	0.01	3.67	< .001***	.07
	Emotion: happy	0.02	0.01	3.01	< .01**	
	Depression group: with depression	0.05	0.07	0.77	.44	.01
	Emotion: angry x Depression group: with depression	-0.02	0.01	-1.80	.07	.02
	Emotion: happy x Depression group: with depression	-0.01	0.01	-1.31	.19	
Normalized intensity range	Intercept	0.97	0.06	17.12	< .001***	
	Emotion: angry	0.21	0.01	16.87	< .001***	.79
	Emotion: happy	0.14	0.01	10.88	< .001***	
	Depression group: with depression	0.06	0.08	0.69	.49	.01
	Emotion: angry x Depression group: with depression	0.03	0.02	1.61	.11	.02
	Emotion: happy x Depression group: with depression	0.01	0.02	0.69	.49	

Table 1 (continued)

Acoustic cue	Parameter	Estimate	*SE*	*t*	*p*	η^2
Speech rate	Intercept	3.57	0.07	53.04	< .001***	
	Emotion: angry	-0.36	0.05	-6.95	< .001***	.47
	Emotion: happy	-0.27	0.05	-5.20	< .001***	
	Depression group: with depression	-0.03	0.09	-0.36	.72	.03
	Emotion: angry x Depression group: with depression	-0.11	0.08	-1.48	.14	.04
	Emotion: happy x Depression group: with depression	-0.19	0.08	-2.44	.02*	

SE = standard error. Neutral is the reference category for emotion type; mothers without a history of depression are the reference category for depression group. Effects of sentence were significant for all acoustic cues, *p*s < .001; parameter estimates (for 20 levels of sentence) are omitted for brevity, but are available from the senior author upon request. ***: $p < .001$, **: $p < .01$, *: $p < .05$. Partial eta squared values (η^2) represent computed effect sizes for fixed effects, based on a Type III analysis of variable table derived with Satterthwaite's method (using *anova* and *F_to_eta2* functions in R)

Table 2 Acoustic characteristics of recordings for each emotion type

Emotion type	Normalized pitch mean (Hz)	Normalized pitch range (Hz)	Normalized intensity mean (dB)	Normalized intensity range (dB)	Speech rate (syllables/second)
Happiness	1.25 (0.03)	2.14 (0.05)	0.86 (0.04)	1.16 (0.04)	3.07 (0.05)
Anger	0.90 (0.03)	1.51 (0.05)	0.86 (0.04)	1.25 (0.04)	3.02 (0.05)
Neutral	0.47 (0.03)	1.07 (0.05)	0.85 (0.04)	1.02 (0.04)	3.44 (0.05)

Values represent estimated marginal means (standard errors of the mean). Hz = Hertz, dB = decibels

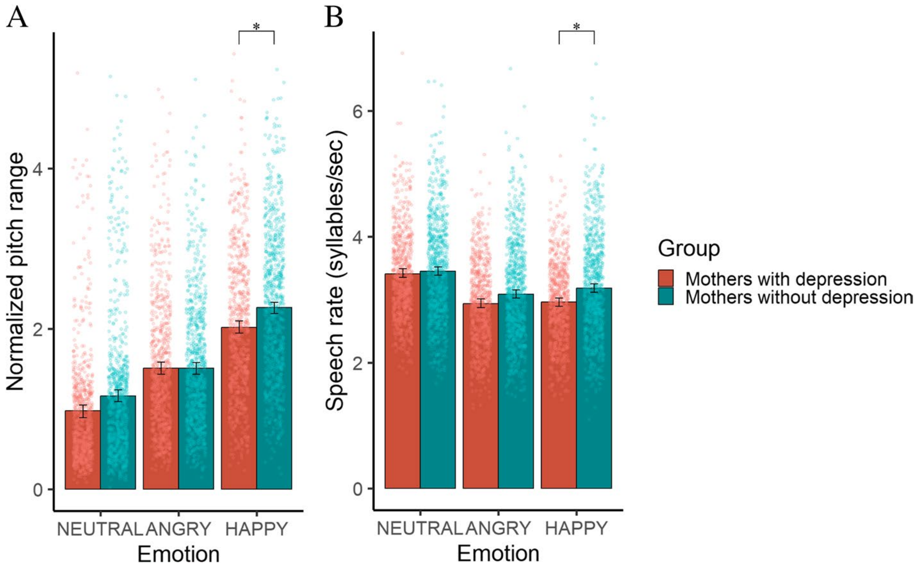


Fig. 1 Interaction of emotion type and depression group on normalized pitch range and speech rate. *Note:* Graphs represent the interaction of emotion type and depression group on normalized pitch range (**A**) and speech rate (**B**). Bars represent mean values (with standard errors of the mean) for mothers with and without a history of depression; each dot represents a recording. * $p < .05$

higher than neutral ($ps < 0.001$). Neutral was lower in intensity mean, but faster in speech rate, than both anger ($ps < 0.004$) and happiness ($ps < 0.02$), which did not differ from one another. Descriptive statistics are provided in Table 2.

There were no main effects of depression group on the acoustic characteristics of mothers' voices in the vocal expression task (i.e., no group difference in overall pitch mean, pitch range, mean intensity, intensity range, or speech rate) when collapsed across emotion type, $ps > 0.08$. However, there was a significant interaction of emotion type and depression group for pitch range, $\chi^2(2) = 6.84$, $p = 0.03$ (Table 1; Fig. 1A): simple-effects tests revealed that the MHD group had a significantly lower pitch range ($M = 2.02$, $SE = 0.08$) than the MND group ($M = 2.26$, $SE = 0.07$) when expressing happiness, $p = 0.02$. There was also a significant interaction between emotion type and depression group for speech rate, $\chi^2(2) = 5.93$, $p = 0.05$ (Table 1; Fig. 1B): simple-effects tests indicated that MHD spoke more slowly ($M = 2.96$, $SE = 0.07$) when expressing happiness, compared to MND ($M = 3.18$, $SE = 0.06$), $p = 0.02$.

Table 3 Descriptive statistics for listeners' perceptual ratings of recordings, by depression group

Perceptual rating	Mothers with depression	Mothers without depression
Accuracy	0.63 (0.02)	0.64 (0.02)
Emotional intensity	2.48 (0.06)	2.50 (0.06)
Recognizability	2.51 (0.07)	2.53 (0.07)
Authenticity	2.33 (0.05)	2.35 (0.05)

Values represent listeners' mean ratings (standard error of the mean) for mothers with and without a history of depression. Accuracy ranges from 0 (incorrect) to 1 (correct); emotional intensity, recognizability, and authenticity ratings range from 1 to 4

Listeners' Perceptions of Emotional Recordings

Listeners' ratings suggested that mothers' angry and happy recordings were well-recognized as conveying their intended emotion type. Overall accuracy was 63.64%. One-sample *t*-tests indicated that accuracy in emotion identification was above chance level (i.e., 33%) for each emotion type (76.97% accuracy for anger, 50.29% accuracy for happiness), $ps < 0.001$. Ratings were significantly above 2 on the 1–4 scale for emotional intensity, $p < 0.001$, recognizability, $p < 0.001$, and authenticity, $p < 0.001$, in one-sample *t*-tests (descriptive statistics provided in Table 3).

Results of the linear mixed-effect models indicated that recordings produced by mothers with and without a history of depression did *not* differ in recognition accuracy (i.e., listeners were not more or less able to identify the intended emotion), $p = 0.91$, nor in ratings of their emotional intensity, $p = 0.84$, recognizability, $p = 0.79$, or authenticity, $p = 0.79$. This effect did not vary across emotion types ($ps > 0.10$), except for a significant interaction between emotion type and depression group on authenticity, $\chi^2(2) = 678.58$, $p < 0.001$. However, post-hoc analyses did not indicate any significant differences in ratings between groups, $ps > 0.90$ (anger $M = 2.54$ for MHD, $M = 2.52$ for MND; happiness $M = 2.12$ for MHD, $M = 2.17$ for MND). As such, we do not interpret this interaction further. Parameter estimates (e.g., B, *t*) for all models are presented in Table 4.[6]

Perceptual Correlates of Acoustic Characteristics

Because we found that depression groups differed in the acoustic cues associated with their portrayals of happiness, we conducted secondary analyses to examine a) how happy recordings' acoustic characteristics influenced listeners' perception of them, and b) whether that relationship varied as a function of speakers' depression group. We fit linear mixed-effects models testing the effect of each happy recording's pitch range

[6] Including current BDI scores as a covariate did not alter the findings presented in text. However, BDI scores and depression group membership were highly correlated (Spearman's $\rho = .65$, $p < .001$). To avoid issues with multicollinearity, we therefore fit alternative models for the analyses in Sections "Listeners' perceptions of emotional recordings" and "Perceptual correlates of acoustic characteristics", in which BDI scores were used as a predictor instead of depression group. The effects of emotion type on acoustic cues remain significant; however, there are no main effects of speakers' depression symptom scores nor interactions with emotion type (all $ps > .07$) on acoustic cues. Similarly, speakers' BDI scores were not associated with listeners' accuracy or ratings of their recordings (all $ps > .38$).

Table 4 Parameter estimates for models examining the effects of speakers' depression group and emotion type on listeners' recognition accuracy and perceptual ratings

Perceptual rating	Parameter	Estimate	*SE*	*t*	*p*
Accuracy	Intercept	0.51	0.02	21.82	< .001***
	Depression group: with depression	< −0.01	0.03	-0.22	.83
	Emotion: anger	0.26	< 0.01	27.11	< .001***
	Depression group: with depression x Emotion: anger	< .01	0.01	0.51	.61
Emotional intensity	Intercept	2.38	0.07	36.49	< .001***
	Depression group: with depression	−0.04	0.09	−0.46	.65
	Emotion: anger	0.24	0.02	11.92	< .001***
	Depression group: with depression x Emotion: anger	0.05	0.03	1.65	.10
Recognizability	Intercept	2.38	0.07	36.03	< .001***
	Depression group: with depression	-0.03	0.09	-0.29	.77
	Emotion: anger	0.30	0.02	14.27	< .001***
	Depression group: with depression x Emotion: anger	< .01	0.03	0.20	.84
Authenticity	Intercept	2.17	0.05	42.01	< .001***
	Depression group: with depression	−0.05	0.07	−0.69	.49
	Emotion: anger	0.35	0.02	17.15	< .001***
	Depression group: with depression x Emotion: anger	0.06	0.03	2.06	.04 *

SE = standard error. Mothers without a history of depression are the reference category for depression group; happiness is the reference category for emotion type. ***: $p < .001$, **: $p < .01$

or speech rate (continuous predictors; separate model for each acoustic cue) and its speaker's depression group (categorical predictor) on listeners' ratings of its emotional intensity, recognizability, and authenticity, including both per-speaker and per-listener random intercepts. Models were built by first adding the fixed effect of the acoustic cue to a base model that only included the fixed and random intercepts; we then added the fixed effect of depression group and the interaction term (depression group x acoustic cue) in sequence. Estimated *p* values (reported in text) for all fixed effects were obtained from likelihood ratio tests of the model with the effect in question against the previous model in this model-building sequence. Parameter estimates (e.g., B, *t*) from full models are reported in Table 5.

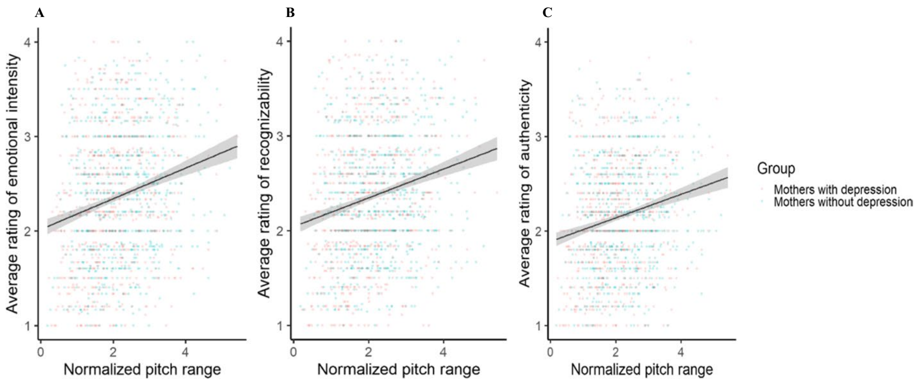
Across both depression groups, happy recordings with lower pitch range were rated as less emotionally intense, $\chi^2(1) = 29.32$, $p < 0.001$, less recognizably happy, $\chi^2(1) = 31.00$, $p < 0.001$, and less authentically happy, $\chi^2(1) = 30.08$, $p < 0.001$ (Fig. 2). However, these effects were of small magnitude (Table 5). There were no main effects of depression group ($ps < 0.61$), nor interactions of depression group and pitch range ($ps < 0.14$), on perceptual ratings.

Speech rate was significantly negatively associated with ratings of happy recordings' emotional intensity, $\chi^2(1) = 55.62$, $p < 0.001$, recognizability, $\chi^2(1) = 54.76$, $p < 0.001$, and authenticity, $\chi^2(1) = 36.85$, $p < 0.001$ (Fig. 3). As above, these effects were small in magnitude (Table 5), and did not differ across depression groups, $ps > 0.27$. However, there was an interaction of depression and speech rate on recognizability ratings,

Table 5 Parameter estimates for models examining the effects of acoustic cues on listeners' perceptual ratings

Perceptual rating	Acoustic cue	Parameter	Estimate	*SE*	*t*	*p*
Emotional intensity	Normalized pitch range	Intercept	2.17	0.08	27.66	< .001***
		Normalized pitch range	0.10	0.02	4.88	< .001***
		Depression group: with depression	0.06	0.11	0.58	.56
		Normalized pitch range x Depression group: with depression	−0.04	0.03	−1.46	.14
	Speech rate	Intercept	2.72	0.10	28.31	< .001***
		Speech rate	−0.11	0.02	−4.70	< .001***
		Depression group: with depression	0.04	0.13	0.33	.75
		Speech rate x Depression group: with depression	−0.04	0.03	−1.15	.25
Recognizability	Normalized pitch range	Intercept	2.16	0.08	27.05	< .001***
		Normalized pitch range	0.10	0.02	4.77	< .001***
		Depression group: with depression	0.06	0.11	0.56	.57
		Normalized pitch range x Depression group: with depression	−0.03	0.03	−1.15	.25
	Speech rate	Intercept	2.69	0.10	27.21	< .001***
		Speech rate	−0.10	0.02	−4.06	< .001***
		Depression group: with depression	0.15	0.14	1.12	.27
		Speech rate x Depression group: with depression	−0.07	0.03	−2.02	.04*
Authenticity	Normalized pitch range	Intercept	1.98	0.07	29.63	< .001***
		Normalized pitch range	0.09	0.02	4.42	< .001***
		Depression group: with depression	0.01	0.09	0.14	.89
		Normalized pitch range x Depression group: with depression	−0.02	0.03	−0.75	.46
	Speech rate	Intercept	2.44	0.09	27.98	< .001***
		Speech rate	−0.08	0.02	−3.80	< .001***
		Depression group: with depression	0.03	0.12	0.24	.81
		Speech rate x Depression group: with depression	−0.03	0.03	−1.03	.30

SE = standard error. Mothers without a history of depression are the reference category for depression group. ***: $p < .001$, **: $p < .01$



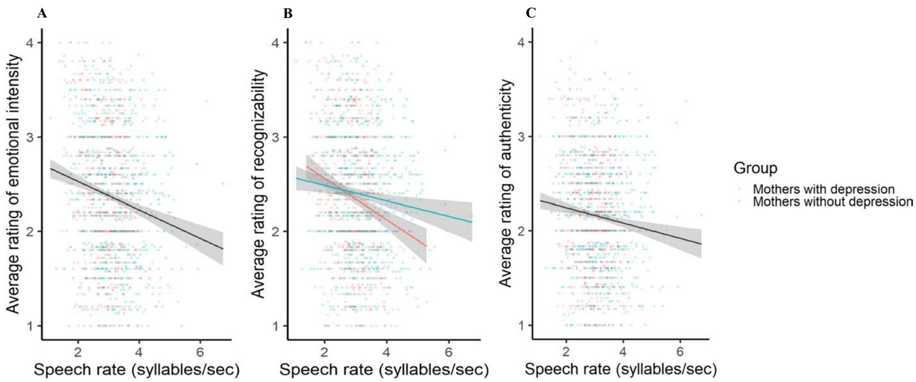
Note: Graphs represent the relationship between normalized pitch range in happy recordings and listeners' average ratings of their emotional intensity (A), recognizability (B), and authenticity (C). Each dot represents a recording. Average ratings are depicted for visualization purposes only; individual listeners' ratings were entered into analyses without being averaged first. Depression groups are defined based on mothers' history of depression.

Fig. 2 Relationship between normalized pitch range and perceptual ratings of happy recordings. *Note:* Graphs represent the relationship between normalized pitch range in happy recordings and listeners' average ratings of their emotional intensity (**A**), recognizability (**B**), and authenticity (**C**). Each dot represents a recording. Average ratings are depicted for visualization purposes only; individual listeners' ratings were entered into analyses without being averaged first. Depression groups are defined based on mothers' history of depression

$\chi^2(1) = 4.07$, $p = 0.04$, whereby the association between speech rate and recognizability was more strongly negative in MHD compared to MND. There was no interaction effect for ratings of emotional intensity ($p = 0.25$), nor for ratings of authenticity ($p = 0.31$).

Discussion

The present study examined whether mothers with (MHD) and without (MND) a history of major depression during their preschool-aged children's lifetime differed in their vocal expressions of happiness and anger. Mothers were asked to express child-relevant sentences in neutral, happy, and angry tones of voice. We examined whether MHD and MND expressed vocal emotions using different acoustic cues (e.g., pitch, intensity, speech rate). We also asked a large group of independent listeners to rate recordings on intended emotion (accuracy), as well as the intensity, recognizability and authenticity of each expression. We found that MHD expressed happiness with less range in pitch and a slower speech rate (i.e., with a slower, more monotonous voice) than MND. However, listeners' ratings did not differ between groups. Across both groups, exemplars of happiness that were expressed with lower pitch range were perceived as less intensely, recognizably, and authentically happy by listeners. Interestingly, slower speech rate was associated with the opposite pattern—a finding that was unexpected given previous work on the typical acoustic characteristics of happy expressions (Johnstone & Scherer, 2000), but that may be reflective of the child-directed nature of our speech sentences. Taken together, our findings indicate that MHD may express happiness in acoustically atypical ways, but that their vocal expressions may not necessarily be uniformly harder to recognize or perceived as less authentic/recognizable/intense by adult listeners. Instead, idiosyncratic vocal patterns, such as blunted pitch range when expressing happiness, may be selectively difficult for listeners to interpret (at an individual speaker level, rather than a group level). Given that vocal cues associated with emotional expressions are generally under some level of voluntary control, our



Note: Graphs represent the relationship between speech rate in happy recordings and listeners' average ratings of their emotional intensity (A), recognizability (B), and authenticity (C). Graph B represents the significant interaction of depression group x speech rate on recognizability ratings. Each dot represents a recording. Average ratings are depicted for visualization purposes only; individual listeners' ratings were entered into analyses without being averaged first. Depression groups are defined based on mothers' history of depression.

Fig. 3 Relationship between speech rate and perceptual ratings of happy recordings. *Note:* Graphs represent the relationship between speech rate in happy recordings and listeners' average ratings of their emotional intensity (**A**), recognizability (**B**), and authenticity (**C**). Graph B represents the significant interaction of depression group x speech rate on recognizability ratings. Each dot represents a recording. Average ratings are depicted for visualization purposes only; individual listeners' ratings were entered into analyses without being averaged first. Depression groups are defined based on mothers' history of depression

results highlight potential ways in which parents may improve the clarity of their emotional message in interactions with their children (e.g., by purposefully varying pitch range in a more pronounced way when attempting to convey positive affect). This may be particularly relevant to parents with a history of depression, who as a group tended to exhibit this pattern. Our findings also highlight future directions for research on the impact of parents' emotional expression styles on children's emotional development.

Mothers' Vocal Expression of Emotion

Complementing previous work finding that people with depression exhibit differences in spontaneous vocal prosody compared to those without depression (e.g., Alpert et al., 2001; Cannizzaro et al., 2004; Darby et al., 1984; McGinnis et al., 2019; Murray et al., 2010), our task elicited the purposeful encoding of *emotional* prosody. We found emotion-specific group differences in pitch range and speech rate: MHD used less variation in pitch and a slower speech rate when enacting happiness than did MND. These effects were small in magnitude, but group differences exceeded reported just-noticeable differences in pitch and speech tempo (4.1% and 5% respectively; Re et al., 2012; Quené, 2007)—meaning that they were likely above perceptually detectable thresholds. These patterns are contrary to those typically associated with the expression of happiness, which is generally conveyed using a high-pitched voice, high variation in pitch (i.e., high pitch range), and a fast speech rate (Johnstone & Scherer, 2000). These prosodic components of happiness were diminished in MHD, suggesting that these mothers may be enacting acoustically atypical exemplars of happiness.

Such differences may be a consequence of the types of emotional experiences typically associated with depression. Anhedonia, a major component of depression (Pizzagalli, 2014), could impact the ability to infuse speech with positive and happy tones. In addition, individuals with depression can be more prone to negative emotionality, either towards themselves or other people (Busch, 2009). When talking to their infants, mothers with depression have been found to have more acoustic elements typical of sadness—which is associated with a slower rate of speech (Johnstone & Scherer, 2000)—in their speech (Murray et al., 2010). Therefore, it is possible that MHD had a more challenging time overcompensating for the elements of sadness or slowed speech in their natural way of speaking when trying to enact a positive valence emotion (like happiness) in the current vocal expression task. Furthermore, individuals with depression tend to have more of a "flat expression" in their voices (Alpert et al., 2001), as well as a flattening of positive affect in general (Bean et al., 2022), which may stem from blunted responsivity to positive stimuli (Bylsma et al., 2008). Based on these prior findings, we speculate that mothers with depression may be less inclined to produce the "up and down" pitch and faster speech rate associated with prototypically happy prosody, as it may be more distinct from their natural patterns of speech. Thus, the symptomatology of depression—including greater likelihood of experiencing negative affect, reduced positive emotion (Bylsma et al., 2008), and blunted nonverbal expressions—could be contributing to less prototypical expressions of happiness in MHD. In this context, it is noteworthy that group differences in pitch range and speech rate were specific to the expression of happiness. The fact that no differences were noted in angry prosody suggests that effects of depression history on nonverbal expression patterns may be emotion- or valence-specific. Future studies could disentangle the source of variance in these effects by testing a wider group of emotion types, including sadness.

Perception of Mothers' Vocal Expressions by Listeners

Overall, mothers' portrayals of vocal emotion were consistent with established acoustic patterns of happiness and anger (see review by Johnstone & Scherer, 2000); listeners' recognition of the intended emotion in these portrayals was also on par with expected levels of accuracy for vocal prosody (i.e., approximately 60%; see review by Scherer, 2003). Contrary to our hypotheses, we did not find that the vocal expressions of MHD as a whole were perceived as less emotionally intense, recognizable, or authentic than those of their counterparts without depression histories. It is possible that the acoustic cues that differentiated both depression groups were not pronounced enough to impact the perception of these expressions at a group level. Although alterations in pitch range and speech rate do influence listeners' impressions of emotionality in prosody (Scherer & Oshinsky, 1977), other acoustic cues like mean pitch—which did not vary by depression group—account for larger proportions of variance in listeners' interpretations of emotional prosody (Scherer, 1996). As such, the *ensemble* of acoustic cues contained in happy and angry expressions produced by MHD may be yielding typical perceptual ratings of their vocal prosody, despite their differential deployment of specific vocal cues.

Nonetheless, across groups, exemplars of happiness characterized by lower range in pitch—a pattern more likely to occur in mothers with MHD—were rated as less emotionally intense, less recognizable, and less authentically happy. Since happiness is typically expressed with *greater* range in pitch (Johnstone & Scherer, 2000), happy recordings with lower pitch range were unsurprisingly perceived as poorer quality exemplars of happiness. Typically, speaking faster (i.e., with a higher speech rate) is also associated with the expression of happiness (Johnstone & Scherer, 2000). However, in our sample, happy exemplars expressed with a *lower* speech rate were rated as sounding more intensely, recognizably, and authentically happy (with this association particularly pronounced for recognizability in the depression group). Although both effects were small, this surprising pattern may be reflective of task demands: in the current vocal expression task, mothers were instructed to portray their emotions using child-directed sentences (e.g., "Put your shoes on", or "It's time to get dressed!"). This may have pulled for child-directed speech (or 'motherese'), a type of speech pattern commonly expressed by parents towards their young children. Child-directed speech is typically characterized by an elongation of utterances (particularly the last vowel) and a sing-song quality, which would translate to a slowed speech rate and a higher pitch range (Grieser & Kuhl, 1988; Morningstar et al., 2019; Saint-Georges et al., 2013). In our study, we observed that the MHD group were slowing their speech when expressing happiness to a greater extent than the MND group, but used a blunted variation in pitch (i.e., less of a sing-song quality; more monotone). Mothers with depression have been found to use less of the exaggerated prosody characteristic of infant-directed speech (and less child-directed speech in general; Scheiber et al., 2022), which was linked to impaired infant associative learning in an attention task (Kaplan et al., 2002; Lam-Cassettari & Kohlhoff, 2020). Mothers diagnosed with and partially remitted from clinical depression have been found to use less range in pitch during play interactions with their infants, compared to non-depressed or fully remitted mothers (Porritt et al., 2014). Since the use of child-directed speech is thought to promote language acquisition and social interaction with parents (Golinkoff et al., 2015), atypical variations in its enactment may have downstream effects on learning or emotion socialization processes for children.

Because the current task did not assess affective prosody in naturalistic interactions between mother and child, whether atypical expressions of posed happiness would also be detectable in dyadic interactions and have similar effects on children is an open question. Children often use their parents as baselines to understand emotion (Halberstadt, 1991); as such, having a parent whose presentation of positive affect is less clearly communicated could affect a child's reference point for positive emotional expression. To investigate this possibility, future studies could ask children to rate these recordings to see whether offspring of mothers with depression perceive these vocal expressions differently than children who do not have a mother with depression (as with facial expressions: Joormann et al., 2010; Székely et al., 2014). Examining parental expressions and child perceptions in a longitudinal framework would also offer insight into whether exposure to atypical vocal enactments of happiness impacts children's understanding of emotions, or plays a role in the intergenerational transmission of depression (Goodman, 2020). Futhermore, it is also possible that differences in emotional expression styles may have impacts in other domains of development, such as attention, (social) communication, and processing of social reward. Future research should examine these questions further.

Strengths and Limitations

Whereas previous studies have focused on the effect of depression status on characteristics of free and unrestrained speech in conversational contexts, our study builds upon this work by examining the impact of depression history on the enactment of purposefully emotional speech. By prompting mothers to express different emotional states with their voices, we were better able to investigate the association between depression and *de novo* displays of emotion. Furthermore, by controlling for speech content in an experimental setting, our experimental design prevented participants from choosing emotionally-valenced topics, which could impact affective prosody in unsystematic ways (i.e., choosing to talk about a frustrating scenario could impact the prosodic cues one speaks with). Although we are not capturing spontaneous emotional cues, social information processing theories posit that the ability to convey one's own emotional state to others is crucial to successful social interactions (Crick & Dodge, 1994; Halberstadt et al., 2001; Lemerise

& Arsenio, 2000). The current study thus provides novel information about the impact of depression on an important facet of socio-emotional communication.

One limitation of the current study is that the inclusion criteria for the depression group was not limited to a current major depressive episode. Although mothers in the MHD group did report more current depression symptoms at the time of testing, the majority of mothers in this group were not currently experiencing a major depressive episode. As such, our inclusion criterion may have introduced heterogeneity within the MHD group. In addition, our study is limited by the fact that we did not examine how group differences in maternal prosodic expression relate to symptom expression in childen. As internalizing symptoms in children tend to appear later in development, such as during adolescence (Kessler et al., 2005), it may be beneficial to assess how subsequent periods of development for the child may be impacted by maternal prosodic differences in the context of depression.

Additonally, we did not have information about mothers' previous acting experience, which could have affected both their portrayals and listeners' judgments of the quality of their emotional expressions. Further, happiness was always portrayed after anger in the vocal recording task; while this was done to neutralize any negative emotion induction by the end of the testing session (something that was deemed important given the population of interest), this represents a confound between emotion type and condition order. We incorporated breaks between emotion types to minimize cross-over effects; however, we cannot entirely rule out the possibility of order effects (e.g., that MHD may have struggled to upregulate the purposeful expression of emotion across two conditions, rather than specifically deviated in their portrayal of happiness).

Future studies should also examine how depression impacts multiple nonverbal modalities in combination. As stated previously, depression not only affects vocal prosody but a myriad of nonverbal cues such as facial expressions (Cohn et al., 2009; Gehricke & Shapiro, 2000; Girard et al., 2013) and postures (Feldman et al., 2020; Michalak et al., 2009). Although vocal cues are an important component of dyadic communication—including child-directed speech, which plays a large role in the parent–child relationship from an early age (Saint-Georges et al., 2013)—examining nonverbal cues additively or multimodally would provide a more complete picture of the impact of depression history on maternal emotional communication.

Lastly, our sample consisted of primarily white mother–child dyads and was skewed toward higher SES representation. As such, our findings may not be generalizable to lower income or more diverse populations. Our sample of raters were also college students who primarily identified as women, which may further limit generalizability of our results. Examining these questions in a more diverse group of speakers and listeners will be important.

Conclusion

The current study examined whether mothers with and without a history of depression varied in their posed expression of vocal emotions, both objectively (based on speech analysis) and perceptually (based on listener ratings). Results indicated that MHD expressed happiness with less pitch range and slower speech (i.e., a more monotone voice) than MND. Furthermore, we found that the acoustic properties of their emotional expressions had implications for the way that listeners rated the recordings: across groups, happy expressions with lower pitch range were rated as less emotionally intense, authentic and recognizably happy, although faster speech rate was associated with the opposite pattern. These patterns may be reflective of task demands that elicited child-directed speech: MHD may have been enacting the slowed speech rate of child-directed speech without the associated variation in pitch. Taken together, our findings suggest that acoustic variations in how MHD express vocal emotions may not directly translate to differential perception of their expressions by adult listeners at a group level; however, the atypical combination of acoustic cues is likely to complicate the unambiguous communication of emotional intent.

These results add to the existing literature in a novel way by delineating an association between a history of depression and *de novo* emotional prosody. Future research should investigate the relationship between acoustic differences in emotional expression and maternal behaviour with their child. We identified differences in the production of emotional prosody at a group level; additional work needs to determine what the perceptual consequences may be at a parent–child dyad level. Vocal cues, including pitch range, are a modifiable aspect of socio-emotional communication. For instance, parents who completed a parent–child interaction intervention were found to spontaneously modulate this aspect of their voice in subsequent interactions with their child (Morningstar et al., 2019). It may thus be possible for parents with a history of depression—who may be more likely to express happiness with an atypical combination of vocal cues—to improve the clarity of their emotional messages by purposefully varying their pitch when expressing happiness. Future studies would benefit from tracking changes in parental vocal expression of emotion—perhaps as a function of remission from depression—in parent–child dyads to better understand the functional consequences of nonverbal communication on children's understanding of emotion.

Appendix

1	Why did you do that?
2	When will you be done?
3	Where did you get that?
4	What do you need?
5	Are you finished yet?
6	Did you hear me?
7	Who told you that?
8	Where were you?
9	Did you put your shoes away?
10	Did you do this?
11	I can't believe you just did that!
12	I didn't know about it
13	You shouldn't have done that
14	Bring that with you
15	You should go outside
16	Look at that!
17	Put your seatbelt on
18	It's time to get dressed!
19	Don't forget about it
20	Let's hurry!

Sentences were constructed to be neutral in content, but could feasibly be said in a positive or negative valence.

Acknowledgements We are grateful to Connor Grannis, Andy Hung, Roberto French, Meika Travis, Scout Crowell, Peyton Thomas, Reegan Pomponio, Julie Shin, Stephanie Cline, Madeleine Helesic, Gabrielle Craddock, Rafaela Platkin, Bruno Carusi, and Maya Morcos for their help in collecting and processing the data. We also wish to thank the participants and their families for their time.

Author contributions E.I. and M.M. wrote the main manuscript text, prepared all tables and figures, and conceptualized the study's research questions. E.I. and M.M. conducted all analyses in the study, with assistance from X.F and X.F. All authors were involved in study design and the interpretation of findings. All authors reviewed and provided edits to the manuscript.

Funding This work was supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant to M. Morningstar; GPIN-2021–03130), the Social Sciences and Humanities Research Council of Canada (Canada Graduate Scholarship to E. Ilyaz), Eunice Kennedy Shriver National Institute of Child Health and Human Development (R21 HD095179 awarded to X. Feng), and the Abigail Wexner Research Institute at Nationwide Children's Hospital (intramural funds to E. Nelson).

Data Availability Listeners' data is available upon request to the corresponding author.

Declarations

Ethical Approval All procedures performed in studies involving human participants were in accordance with the 1964 Declaration of Helsinki and its later amendments.

Competing Interests The authors declare no competing interests.

References

Alpert, M., Pouget, E. R., & Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders, 66*(1), 59–69. https://doi.org/10.1016/S0165-0327(00)00335-9

Audacity Team. (2020). Audacity: Free audio editor and recorder [Computer program]. Version 2.4.2 retrieved from https://www.audacityteam.org/download/

Azak, S., & Raeder, S. (2013). Trajectories of parenting behavior and maternal depression. *Infant Behavior and Development, 36*(3), 391–402. https://doi.org/10.1016/j.infbeh.2013.03.004

Bachorowski, J. A. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science, 8*(2), 53–57.

Barrett, K. C., & Campos, J. J. (1987). Perspectives on emotional development II: A functionalist approach to emotions. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 555–578). John Wiley & Sons.

Bean, C. A., Summers, C. B., & Ciesla, J. A. (2022). Dampening of positive affect and depression: A meta-analysis of cross-sectional and longitudinal relationships. *Behaviour Research and Therapy, 156*, 104153. https://doi.org/10.1016/j.brat.2022.104153

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II : Psychological Corporation.* https://doi.org/10.1037/t00742-000

Boersma, P., & Weenick, D. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9–10), 341–345.

Busch, F. N. (2009). Anger and depression. *Advances in Psychiatric Treatment, 15*(4), 271–278. https://doi.org/10.1192/apt.bp.107.004937

Bylsma, L. M., Morris, B. H., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review, 28*(4), 676–691. https://doi.org/10.1016/j.cpr.2007.10.001

Cannizzaro, M., Harel, B., Reilly, N., Chappell, P., & Snyder, P. J. (2004). Voice acoustical measurement of the severity of major depression. *Brain and Cognition, 56*(1), 30–35. https://doi.org/10.1016/j.bandc.2004.05.003

Capatides, J. B., & Bloom, L. (1993). Underlying process in the socialization of emotion. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research.* Erlbaum.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*(3), 316–336. https://doi.org/10.1037/0021-843X.100.3.316

Cohn, J. F., Kruez, T. S., Matthews, I., Yang, Y., Nguyen, M. H., Padilla, M. T., ..., & De la Torre, F. (2009). Detecting depression from facial actions and vocal prosody. Paper presented at the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, Amsterdam, Netherlands. Abstract retrieved from https://doi.org/10.1109/ACII.2009.5349358

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74. https://doi.org/10.1037/0033-2909.115.1.74

Darby, J. K., Simmons, N., & Berger, P. A. (1984). Speech and voice parameters of depression: A pilot study. *Journal of Communication Disorders, 17*(2), 75–85. https://doi.org/10.1016/0021-9924(84)90013-3

Feldman, R., Granat, A. D. I., Pariente, C., Kanety, H., Kuint, J., & Gilboa-Schechtman, E. (2009). Maternal depression and anxiety across the postpartum year and infant social engagement, fear regulation, and stress reactivity. *Journal of the American Academy of Child & Adolescent Psychiatry, 48*(9), 919–927. https://doi.org/10.1097/CHI.0b013e3181b21651

Feldman, R., Schreiber, S., Pick, C. G., & Been, E. (2020). Gait, balance and posture in major mental illnesses: Depression, anxiety and schizophrenia. *Austin Medical Sciences, 5*(1), 1–6.

First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2016). *User's guide for the SCID-5-CV Structured Clinical Interview for DSM-5 disorders: Clinical version*. American Psychiatric Publishing, Inc.

Forbes, E. E., Williamson, D. E., Ryan, N. D., & Dahl, R. E. (2004). Positive and negative affect in depression: Influence of sex and puberty. *Annals of the New York Academy of Sciences, 1021*(1), 341–347. https://doi.org/10.1196/annals.1308.042

Gehricke, J.-G., & Shapiro, D. (2000). Reduced facial expression and social context in major depression: Discrepancies between facial muscle activity and self-reported emotion. *Psychiatry Research, 95*(2), 157–167. https://doi.org/10.1016/S0165-1781(00)00168-2

Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S., & Rosenwald, D. P. (2013). Social risk and depression: Evidence from manual and automatic facial expression analysis. *IEEE Transactions on Affective Computing, 4*(2), 142–150. https://doi.org/10.1109/T-AFFC.2012.38

Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S. M., Hammal, Z., & Rosenwald, D. P. (2014). Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses. *Image and Vision Computing, 32*(10), 641–647. https://doi.org/10.1016/j.imavis.2013.12.007

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

Goodman, S. H. (2020). Intergenerational transmission of depression. *Annual Review of Clinical Psychology, 16*, 213–238. https://doi.org/10.1146/annurev-clinpsy-071519-113915

Goy, H., Fernandes, D. N., Pichora-Fuller, M. K., & van Lieshout, P. (2013). Normative voice data for younger and older adults. *Journal of Voice, 27*(5), 545–555.

Gravener, J. A., Rogosch, F. A., Oshri, A., Narayan, A. J., Cicchetti, D., & Toth, S. L. (2012). The relations among maternal depressive disorder, maternal expressed emotion, and toddler behavior problems and attachment. *Journal of Abnormal Child Psychology, 40*, 803–813. https://doi.org/10.1007/s10802-011-9598-z

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24*(1), 14.

Halberstadt, A. G. (1991). Toward an ecology of expressiveness: Family socialization in particular and a model in general. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 106–160). Editions de la Maison des Sciences de l'Homme; Cambridge University Press.

Halberstadt, A. G., Denham, S. A., & Dunsmore, J. C. (2001). *Affective Social Competence. Social Development, 10*(1), 79–119. https://doi.org/10.1111/1467-9507.00150

Hoffman, C., Crnic, K. A., & Baker, J. K. (2006). Maternal depression and parenting: Implications for children's emergent emotion regulation and behavioral functioning. *Parenting: Science and Practice, 6*(4), 271–295. https://doi.org/10.1207/s15327922par0604_1

Hooper, E., Feng, X., Christian, L., & Slesnick, N. (2015). Emotion expression, emotionality, depressive symptoms, and stress: Maternal profiles related to child outcomes. *Journal of Abnormal Child Psychology, 43*, 1319–1331. https://doi.org/10.1007/s10802-015-0019-6

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. *Handbook of Emotions, 2*, 220–235.

Joormann, J., & Stanton, C. H. (2016). Examining emotion regulation in depression: A review and future directions. *Behaviour Research and Therapy, 86*, 35–49. https://doi.org/10.1016/j.brat.2016.07.007

Joormann, J., & Vanderlind, W. M. (2014). Emotion regulation in depression: The role of biased cognition and reduced cognitive control. *Clinical Psychological Science, 2*(4), 402–421. https://doi.org/10.1177/216770261453616

Joormann, J., Gilbert, K., & Gotlib, I. H. (2010). Emotion identification in girls at high risk for depression. *Journal of Child Psychology and Psychiatry, 51*(5), 575–582. https://doi.org/10.1111/j.1469-7610.2009.02175.x

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*(5), 770.

Kaplan, P. S., Bachorowski, J. A., Smoski, M. J., & Hudenko, W. J. (2002). Infants of depressed mothers, although competent learners, fail to learn in response to their own mothers' infant-directed speech. *Psychological Science, 13*(3), 268–271.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593–602. https://doi.org/10.1001/archpsyc.62.6.593

Kolacz, J., daSilva, E. B., Lewis, G. F., Bertenthal, B. I., & Porges, S. W. (2022). Associations between acoustic features of maternal speech and infants' emotion regulation following a social stressor. *Infancy, 27*(1), 135–158. https://doi.org/10.1111/infa.12440

Kujawa, A., & Burkhouse, K. L. (2017). Vulnerability to depression in youth: Advances from affective neuroscience. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2*(1), 28–37. https://doi.org/10.1016/j.bpsc.2016.09.006

Kujawa, A., Dougherty, L. E. A., Durbin, C. E., Laptook, R., Torpey, D., & Klein, D. N. (2014). Emotion recognition in preschool children: Associations with maternal depression and early parenting. *Development and Psychopathology, 26*(1), 159–170. https://doi.org/10.1017/S0954579413000928

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lam-Cassettari, C., & Kohlhoff, J. (2020). Effect of maternal depression on infant-directed speech to prelinguistic infants: Implications for language development. *PLoS ONE, 15*(7), e0236787. https://doi.org/10.1371/journal.pone/0236787

Lee, J. H., & Humes, L. E. (2012). Effect of fundamental-frequency and sentence-onset differences on speech-identification performance of young and older adults in a competing-talker background. *Journal of the Acoustical Society of America, 132*(3), 1700–1717.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development, 71*(1), 107–118. https://doi.org/10.1111/1467-8624.00124

Loechner, J., Sfärlea, A., Starman, K., Oort, F., Thomsen, L. A., Schulte-Koerne, G., & Platt, B. (2020). Risk of depression in the offspring of parents with depression: The role of emotion regulation, cognitive style, parenting and life events. *Child Psychiatry & Human Development, 51*, 294–309. https://doi.org/10.1007/s10578-019-00930-4

McGinnis, E. W., Anderau, S. P., Hruschak, J., Gurchiek, R. D., Lopez-Duran, N. L., Fitzgerald, K., Rosenblum, K. L., Muzik, M., & McGinnis, R. S. (2019). Giving voice to vulnerable children: Machine learning analysis of speech detects anxiety and depression in early childhood. *IEEE Journal of Biomedical and Health Informatics, 23*(6), 2294–2301. https://doi.org/10.1109/JBHI.2019.2913590

Michalak, J., Troje, N. F., Fischer, J., Vollmar, P., Heidenreich, T., & Schulte, D. (2009). Embodiment of sadness and depression—gait patterns associated with dysphoric mood. *Psosomatic Medicine, 71*(5), 580–587. https://doi.org/10.1097/PSY.0b013e3181a2515c

Morgan, J. K., Eckstrand, K. L., Silk, J. S., Olino, T. M., Ladouceur, C. D., & Forbes, E. E. (2022). Maternal response to positive affect moderates the impact of familial risk for depression on ventral striatal response to winning reward in 6-to 8-year-old children. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 7*(8), 824–832. https://doi.org/10.1016/j.bpsc.2021.12.014

Morningstar, M., Dirks, M. A., & Huang, S. (2017). Vocal cues underlying youth and adult portrayals of socio-emotional expressions. *Journal of Nonverbal Behavior, 41*(2), 155–183. https://doi.org/10.1007/s10919-017-0250-7

Morningstar, M., Ly, V., Feldman, L., & Dirks, M. A. (2018). Mid-adolescents' and adults' recognition of vocal cues of emotion and social intent: Differences by expression and speaker age. *Journal of Nonverbal Behavior, 42*, 237–251.

Morningstar, M., Garcia, D., Dirks, M. A., & Bagner, D. M. (2019). Changes in parental prosody mediate effect of parent-training intervention on infant language production. *Journal of Consulting and Clinical Psychology, 87*(3), 313. https://doi.org/10.1037/ccp0000375
Morris, A. S., Silk, J. S., Steinberg, L., Myers, S. S., & Robinson, L. R. (2007). The role of the family context in the development of emotion regulation. *Social Development, 16*(2), 361–388. https://doi.org/10.1111/j.1467-9507.2007.00389.x
Morris, A. S., Criss, M. M., Silk, J. S., & Houltberg, B. J. (2017). The impact of parenting on emotion regulation during childhood and adolescence. *Child Development Perspectives, 11*(4), 233–238. https://doi.org/10.1111/cdep.12238
Murray, L., Marwick, H., & Arteche, A. (2010). Sadness in mothers' 'baby-talk' predicts affective disorder in adolescent offspring. *Infant Behavior and Development, 33*(3), 361–364. https://doi.org/10.1016/j.infbeh.2010.03.009
O'Connor, E. E., Langer, D. A., & Tompson, M. C. (2017). Maternal depression and youth internalizing and externalizing symptomatology: Severity and chronicity of past maternal depression and current maternal depressive symptoms. *Journal of Abnormal Child Psychology, 45*, 557–568. https://doi.org/10.1007/s10802-016-0185-1
Pan, W., Flint, J., Shenhav, L., Liu, T., Liu, M., Hu, B., & Zhu, T. (2019). Re-examining the robustness of voice features in predicting depression, compared with baseline of confounders. *PLoS ONE, 14*(6), e0218172. https://doi.org/10.1371/journal.pone.0218172
Pell, M. D., Paulmann, S., Dara, C., Alasseri, A., & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics, 37*(4), 417–435. https://doi.org/10.1016/j.wocn.2009.07.005
Pierce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-01193-y
Pizzagalli, D. A. (2014). Depression, stress, and anhedonia: Toward a synthesis and integrated model. *Annual Review of Clinical Psychology, 10*, 393–423.
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
Quené, H. (2007). On the just noticeable difference for tempo in speech. *Journal of Phonetics, 35*(3), 353–362.
R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
Re, D. E., O'Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for very low and very high voice pitch in humans. *PLoS ONE, 7*(3), e32719.
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.
Scheiber, F. A., Ryckman, K. K., & Demir-Lira, O. E. (2022). Maternal depressive symptoms and maternal child-directed speech: A systematic review. *Journal of Affective Disorders, 297*, 194–207. https://doi.org/10.1016/j.jad.2021.10.015
Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*, 227–256.
Scherer, K. R., & Ellgring, H. (2007). Multimodal expression of emotion: Affect programs or componential appraisal patterns? *Emotion, 7*(1), 158–171.
Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion, 1*, 331–346. https://doi.org/10.1007/BF00992539
Scherer, K. R. (1996). Adding the affective dimension: A new look in speech analysis and synthesis. *Proceedings of the 4th International Conference of Spoken Language Processing.* Philadelphia, USA.
Silk, J. S., Shaw, D. S., Skuban, E. M., Oland, A. A., & Kovacs, M. (2006). Emotion regulation strategies in offspring of childhood-onset depressed mothers. *Journal of Child Psychology and Psychiatry, 47*(1), 69–78. https://doi.org/10.1111/j.1469-7610.2005.01440.x
Silva, W. J., Lopes, L., Cavalcanti Galdino, M. K., & Almeida, A. A. (2021). Voice acoustic parameters as predictors of depression. *Journal of Voice,* ePublication ahead of print. https://doi.org/10.1016/j.jvoice.2021.06.018
Spackman, M. P., Brown, B. L., & Otto, S. (2009). Do emotions have distinct vocal profiles? A study of idiographic patterns of expression. *Cognition and Emotion, 23*(8), 1565–1588.
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.

Székely, E., Lucassen, N., Tiemeier, H., Bakermans-Kranenburg, M. J., Van Ijzendoorn, M. H., Kok, R., Jaddoe, V. W. V., Hofman, A., Verhulst, F. C., & Herba, C. M. (2014). Maternal depressive symptoms and sensitivity are related to young children's facial expression recognition: The Generation R study. *Development and Psychopathology, 26*(2), 333–345.
Thompson, R. A. (1991). Emotional regulation and emotional development. *Educational Psychology Review, 3*, 269–307. https://doi.org/10.1007/BF01319934
Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science, 11*(3), 188–195.
Werner-Seidler, A., Banks, R., Dunn, B. D., & Moulds, M. L. (2013). An investigation of the relationship between positive affect regulation and depression. *Behaviour Research and Therapy, 51*(1), 46–56. https://doi.org/10.1016/j.brat.2012.11.001
Woods, K. J., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics, 79*, 2064–2072.
Yang, Y., Fairbairn, C., & Cohn, F. J. (2012). Detecting depression severity from vocal prosody. *IEEE Transactions on Affective Computing, 4*(2), 142–150. https://doi.org/10.1109/T-AFFC.2012.38

Authors and Affiliations

Emma Ilyaz[1] · Xin Feng[2] · Xiaoxue Fu[3,4] · Eric E. Nelson[4,5] · Michele Morningstar[1,4,6]

✉ Michele Morningstar
michele.morningstar@queensu.ca

Emma Ilyaz
emma.ilyaz@queensu.ca

Xin Feng
feng.88@osu.edu

Xiaoxue Fu
xiaoxuef@mailbox.sc.edu

Eric E. Nelson
eric.nelson@nationwidechildrens.org

1 Department of Psychology, Queen's University, 62 Arch Street, Kingston, ON K7L 3L3, Canada

2 Department of Human Sciences, The Ohio State University, Columbus, OH, USA

3 Department of Psychology, University of South Carolina, Columbia, SC, USA

4 Center for Biobehavioral Health, Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH, USA

5 Department of Pediatrics, The Ohio State University, Columbus, OH, USA

6 Centre for Neuroscience Studies, Queen's University, Kingston, ON, Canada