

Prosody Perception in Typically Developing School-aged Children

Rose Thomas Kalathottukaren* and Suzanne C. Purdy

Discipline of Speech Science, School of Psychology, The University of Auckland, New Zealand

*Corresponding author: Rose Thomas Kalathottukaren, Discipline of Speech Science, School of Psychology, The University of Auckland, New Zealand, Tel: +14036305398; E-mail: rkal180@aucklanduni.ac.nz

Received date: February 06, 2017; Accepted date: March 07, 2017; Published date: March 08, 2017

Copyright: © 2017 Kalathottukaren RT, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Purpose: To report normative data for prosody perception abilities in typically developing school-aged children.

Method: Four receptive prosody subtests of the Profiling Elements of Prosody in Speech-Communication (PEPS-C) and the Child Paralanguage subtest of Diagnostic Analysis of Non Verbal Accuracy 2 (DANVA 2) were administered to 45 children divided into three age groups, with mean ages 7.84, 10.13, and 11.90 years.

Results: Overall results indicated significant age-related improvements in performance on PEPS-C Chunking and Contrastive Stress Reception subtests. Accuracy for emotion recognition differed significantly across the two levels of emotion intensity for the DANVA 2. High emotion intensity items yielded better accuracy compared to low intensity items. A confusion matrix for the DANVA 2 showed that errors were not randomly distributed; some pairs of emotions were confused with one another more often than others. The lowest perceptual accuracy was observed for fear and sadness.

Conclusions: Normative data for prosody perception abilities in typically developing school aged children were reported using PEPS-C receptive prosody subtests and DANVA 2 Child Paralanguage subtest. The development of receptive prosodic skills mostly occurs between 7 and 9 years. Findings of this study have clinical implications for assessing prosody perception in atypical populations.

Keywords: Prosody; Typically developing; Children; Normative data

Introduction

Prosody serves to convey emotions and attitudes (affective prosody), indicate question-statement contrasts, distinguish word boundaries (grammatical prosody), emphasize new and relevant information and pragmatic aspects (pragmatic prosody) of speech. It is important to know how children understand different prosodic functions during communication at different ages and the degree of variability that might be expected within an age group. Relative to the studies examining production of prosodic contrasts [1-4], less is known about the perception of prosodic functions in children. Prosody is reported as a neglected field of research when compared to other aspects of language [5]. Although prosodic difficulties were reported in various communication disorders there is lack of normative data for prosody perception abilities in typically developing children. Assessment of prosodic skills in clinical settings is currently constrained by the lack of normative comparison sample. The present study examined prosody perception abilities in 7-12 year old typically developing children using the receptive prosody subtests of the Profiling Elements of Prosody in Speech-Communication [6] and the Child Paralanguage subtest of the DANVA 2 (Diagnostic Analysis of Non Verbal Accuracy 2) [7].

PEPS-C includes subtests to assess listener's understanding of sentence type (question vs. statement; 'Turn-end Reception'), speaker's emotion (happy or sad; 'Affect Reception'), phrase boundaries (the distinction between simple and compound nouns and groupings of adjectives; 'Chunking Reception'), and placement of contrastive stress/

accent ('Contrastive Stress Reception'). PEPS-C Affect Reception subtest involves option to assess only two emotions; happy and sad. Hence, DANVA 2 Child paralanguage subtest was used in this study which includes option to assess listener's understanding of four different emotions; happy, sad, angry, and fearful. Moreover, PEPS-C Affect subtest uses single word test items (e.g., names of food items) rather than a sentence context. A positive feature of the DANVA 2 subtest is that it uses sentence level stimuli which are more naturalistic than word level stimuli. Test developers of PEPS-C and DANVA 2 have provided test-retest reliability and internal consistency (Cronbach's alpha) information for these tests (Kalathottukaren, Purdy, Ballard, in press).

Perceptual sensitivity to prosodic cues starts in infancy. Newborns are able to discriminate the rhythm, intonation, and stress patterns of their native language [8,9] and are sensitive to general acoustic properties from very early in development and become attuned to the specific features of prosody of their language by about 9 months [10]. Acoustic analyses of infant cry have reported prosody modulations [11,12]. Studies have reported that 6 month old infants are aware of the typical correlation of syllable lengthening, pitch declination, and pausing that occurs at the boundaries of major linguistic units in English [13-16] and are sensitive to syllable weight and typical pattern of strong-weak syllables that occur in English [17,18]. Jusczyk, Cutler, and Redanz [19] reported that 9 month olds listened longer to lists of stressed-unstressed words (typical of English) than to unstressed-stressed words (atypical of English), suggesting that infants are familiar with the dominant trochaic stress pattern in English.

Development of prosodic contrasts starts early in childhood and matures over time [4,20,21]. Patel and Grigos [3] investigated age-related development in the use of different combinations of acoustic cues (F0, intensity, and duration) to mark question-statement contrast in 4, 7, and 11 year old children. They reported that 4 year olds were unable to reliably use rising F0 contour to signal questions instead used increased final syllable duration, while a combination of F0, intensity, and duration cues were used by 7 year olds. Similar to adult production, the older group relied primarily on F0 changes. This is in line with Patel and Brayton's [2] findings that listeners' accuracy in identifying question-statement contrasts and contrastive stress patterns produced by 4 year olds was significantly poorer than for 7 and 11 year olds, suggesting improved stabilization of prosodic control occurring between 4 and 7 years. These findings are further corroborated by Grigos and Patel's [1] study showing that children as young as 4 years old are able to modify their lip and jaw movements to distinguish between declarative-interrogative contrasts, however refinement of these movements continues between 7 and 11 years.

The functions of prosody have been identified at the grammatical, emotional, and pragmatic levels of communication [22,23]. Prosodic cues such as voice onset time, pitch contour, coarticulation, and syllable duration helps word segmentation in children and adults [24-26]. Research on affective speech has reported that prosodic cues are used to express vocal emotions and attitudes [27,28]. d'Alessandro [29] reported voice quality as one of the prosodic cues related to production and perception of emotions. Use of high pitch at the end of the utterance to signal turn-taking [30,31] and pitch accents to convey "new" and already "given" information are examples of the pragmatic functions of prosody [32,33]. Wells et al. [23] examined perception and production of turn-end (question/statement), affect (like/dislike), chunking (fruit, salad, and milk/fruit-salad and milk), and contrastive stress (BLUE and green socks vs. blue and GREEN socks) in typically developing UK English speaking children (N=120, aged between 5-13 years) using PEPS-C test. They reported that production of prosodic contrast functions is largely established by 5 years, although specific functional contrasts such as contrastive stress continue to develop up to 9 years. They also reported that the ability to discriminate question-statement and like-dislike contrasts were mostly acquired by 8 years, however the ability to understand contrastive stress patterns and chunking continues to develop between 10 and 13 years. This is supported by Grigos and Patel's [33] findings that the articulatory movements to produce sentential stress start to develop between 7 and 11 years and continue to develop throughout adolescence. De Ruiter [32] reported that there are differences between children (5 and 7 year olds) and adults in using pitch accents for conveying new and relevant information; indicating a development trend. Compared to the 7 year olds, 5 year olds made significantly less use of prosodic cues to convey turn-taking; suggesting that children learn the pragmatic functions of prosody only later. This is in line with Potamianos and Narayanan's [34,35] findings that compared to older speakers (11-14 year olds), 8-10 year old children produced more filled pauses in dialogue, which indicates delays in thinking and responding during conversations. These study findings suggest that there are differential pattern of development for different aspects of prosody; certain functional contrasts are mastered later than others. Most of the studies reviewed investigated specific aspects of prosody in different subgroups of children.

Accurate recognition of affective prosody is important from a developmental perspective because auditory signals can capture attention from someone who is not visually attending to the speaker, as

mostly occurs between infants, toddlers and their caregivers. Burnham [36] reported that infants' perception of their mother's facial expressions was facilitated when auditory information was added. Fernald [37] reported that 5 month old infants respond to vocal emotions presented in the absence of facial expressions, but not vice versa. Early affective development is important as this has been reported as setting the stage for future relationship and behavioural development in children [38]. Significant correlations between emotion understanding and theory of mind, verbal abilities [39], and academic achievement [7] have been reported in typically developing children. Previous research on emotion perception in children has mainly focused on recognition of facial expressions. In addition to facial expressions, the prosodic properties of speech also provide a rich source of information about an individual's affective state. In addition to PEPS-C Affect Reception subtest, the present study used the Child Paralinguistic subtest of DANVA 2 to assess perception of affective prosody in typically developing 7-12 year old children.

The difference between typical and atypical populations in recognizing emotions may be less prevalent when emotional expressions are depicted at stronger or greater intensities than when less intense expressions are presented. However, the intensity of emotional expressions has only occasionally been studied as a factor affecting children's recognition of vocal emotions. Mazefsky and Oswald [40] reported that children with high functioning autism were less accurate than children with Asperger's syndrome and typically developing peers in understanding emotions at low intensities than high intensities. Mazefsky [41] reported that lower accuracy on DANVA 2 low intensity tone of voice cues were related to greater social impairment and lower social competence measured using Child Behaviour Checklist [42,43] and Scales of Independent Behaviour-Revised [44]. High emotion intensity facial expressions and tone of voice cues were not related to any of these measures. These findings are consistent with Baum and Nowicki's [45] findings that greater accuracy on DANVA 2 low intensity emotional items, but not high intensity items, was related to better social competence (teacher ratings using Child Behaviour Checklist) in typically developing 2nd to 6th grade children. How well children can understand low emotion intensity items is important given that in everyday settings emotional expressions are often subtle [46]. Studies investigating the ability to recognise subtle vocal emotion cues in children are extremely limited but could be valuable in early detection of impaired emotion processing.

There are differences in acoustic cues used to produce different emotions. For example, high values of F0 are used for anger, fear, and happiness, whereas low values of F0 for sadness and disgust [47]. Largest F0 standard deviations (SD) were reported for happiness, followed by anger, then disgust, and the smallest for sadness and fear. Anger and happiness are produced with high voice intensity, followed by disgust, fear, and sadness [48]. Juslin and Laukka [47] reported the effects of emotion intensity on the acoustic cues; higher values of F0 (SD) for strong rather than weak intensity items, with largest effects for anger and disgust. Similarly, there are differences in voice intensity, speech rate, pause proportion, attack time, and voice quality depending on the level of emotion intensity and emotion category. Juslin and Laukka [49] reported that acoustic cues are used probabilistically and continuously so that cues are not perfectly reliable but have to be combined. They also suggested that the cues are combined in an additive fashion, and there is a certain amount of "cue trading" in emotional expressions. For example, if speakers cannot vary pitch to express anger, they may compensate by varying loudness

a bit more. Luo et al. [50] investigated affective prosody recognition in cochlear implant simulations and reported a trade-off between spectral resolution and periodicity cues when performing a vocal emotion recognition task. In order to accurately understand emotion recognition abilities in atypical and typical populations, a range of different emotions at different levels of intensity need to be examined.

Purpose of the study

The purpose of this study was to report normative data for prosody perception abilities in typically developing school-aged children. In particular, we asked the following questions:

- Is there a developmental effect on prosody perception abilities in typically developing children? If so, are there variations in the developmental pattern for different aspects of prosody in children aged between 7-12 years?
- Are there differences in affective prosody perception abilities in typically developing children based on the level of emotion intensity and emotion category?

Method

Participants

Forty-five typically developing children (21 boys and 24 girls) participated. Participants were selected by age to form three groups: 7-8 year olds (Mage=7.84, SD=0.35, age range: 7.34-8.68 years, n=14), 9-10 year olds (Mage=10.13, SD=0.59, age range: 9.13-10.92 years, n=16), and 11-12 year olds (Mage=11.90, SD=0.49, age range: 11.22-12.93 years, n=15) (Table 1). Informed written consent was obtained from caregivers/parents and participation was voluntary. All children met the inclusion criteria of normal hearing (passed a pure tone and immittance audiometry screening), spoke New Zealand-English as their primary mode of communication, and had no history of speech, language, and/or hearing difficulties as reported by parents. Testing took place either in a quiet room at child's home or in the sound proof booth.

Age group	N	Gender distribution	Age (in decimal years)		
			M	SD	Range
7-8 years	14	6/8	7.83	0.33	7.34-8.68
9-10 years	15	8/8	10.13	0.59	9.13-10.82
11-12 years	16	7/8	11.90	0.49	11.07-12.93

Table 1: Participant characteristics.

Materials

Profiling elements of prosody in speech-communication (PEPS-C)

Four receptive prosody subtests of PEPS-C (Turn-end, Affect, Chunking, and Contrastive Stress Reception) were used. These receptive subtests involve simple binary choices, with low memory and processing demands [51]. The pass criterion is set at 75% by Wells & Peppé [51] to the avoid possibility of chance scoring.

1. **Turn-end Reception:** This subtest assesses the function of prosody in interaction by making use of conversational 'turns' each consisting of a single word. The turns/words are names of food-items and the opposition of tones indicates whether the item is 'read' or 'stated' as opposed to 'offered' or voiced as a question/inquiry.

2. **Affect Reception:** In order to assess the use of prosody to convey affective meaning, PEPS-C uses the distinction between expressing strong liking as opposed to reservation/dislike. The test items used are names of food-items.

3. **Chunking Reception:** Chunking refers to boundary-signalling or prosodic delineation of the utterance into units for grammatical, semantic, or pragmatic purposes. PEPS-C uses the minor phrase boundaries that can be used to distinguish between items in a list. For example, colour combinations (pink and black&green socks vs. pink&black and green socks) or single and compound food-items (fruit, salad, and milk vs. fruit-salad and milk).

4. **Contrastive Stress Reception:** Contrastive stress refers to the speaker's use of phonetic prominence to indicate which word or syllable is most important in an utterance. For example, BLUE and green socks (emphasis on the first colour) vs. blue and GREEN socks (emphasis on the second colour).

The pre-recorded auditory stimuli were presented using a laptop computer through a GENELEC 6010A active portable loudspeaker (placed directly in front of the participant) at a comfortable level in the normal conversational range (65 - 75 dB SPL) measured using a sound level meter at the position of the participant's seat. The computer response screen of the PEPS-C involves a split-screen display of cartoon-type pictures. Participants were instructed to either point to the correct item on the screen or to give a verbal response. Before each task, demonstration items and practice items were played to ensure participants' understanding of the task. The automatic scoring provided the raw scores, percentage scores, standard deviation from the normative mean, and a pass/fail indicator. Details of PEPS-C subtests and instructions for administration and scoring are described in Peppé and McCann [6] and on the PEPS-C website (<http://www.peps-c.com>). Reviews of the strengths and weaknesses of the PEPS-C test are provided in Gibbon and Smyth [52], Peppé [53], and Diehl and Paul [5].

Child paralinguistic subtest of diagnostic analysis of non-verbal accuracy 2 (DANVA 2)

The DANVA 2 test was developed by Baum and Nowicki [45] to measure competence in affect recognition by reading facial expressions and voice tone (affective prosody). It includes five subtests: 1) Child Faces, 2) Adult Faces, 3) Child Paralinguistic, and 4) Adult Paralinguistic, and 5) Child and Adult Posture. The current study used the Child Paralinguistic subtest of DANVA 2 to assess emotion recognition using voice only. This 24-item (4 alternative forced choice response paradigm) subtest involved a sentence "I am going out of the room now but I will be back later" presented in happy, sad, angry, and fearful tones at two levels (high and low) of emotion intensity (12 items per intensity level) by male and female speakers (in random sequence). The auditory stimuli were presented through a loudspeaker (using a similar procedure to the PEPS-C) and participants either gave a verbal response by saying if the person sounded happy, sad, angry, or fearful or pointed to the correct emotional smiley faces showing these emotions (Figure 1). Tables showing the number of errors for each emotion, number of errors for high and low intensity items, number of

errors for emotion by intensity, and the responses that were chosen when there was an error were generated using the DANVA 2 automatic scoring. Error profiles can be used to identify the pattern of difficulty. Additional information about the DANVA 2 test can be found on <http://psychology.emory.edu/clinical/interpersonal/>



Figure 1: Response alternatives (happy, sad, angry, and fearful faces) for DANVA 2 Child Paralinguage subtest. After each stimulus is presented, participants made decisions to choose their responses from one of these emotions.

Statistical analyses

Nonparametric tests were used as the data was not normally distributed. Kruskal-Wallis ANOVA tests [54] were used to examine between group differences on PEPS-C and DANVA 2 subtest scores. Post-hoc Mann Whitney U tests were conducted to investigate significant main effects. Friedman ANOVA was used to determine within group differences in scores across PEPS-C tasks and DANVA 2 emotional categories. Post-hoc analyses using Wilcoxon Signed-Rank tests were conducted to examine significant main effects. A Bonferroni correction factor was applied when multiple post-hoc comparisons were performed. IBM SPSS statistics software package (version 22) was used to perform all the statistical tests reported in this study.

Results

Age group differences on PEPS-C receptive prosody tasks

Table 2 shows the mean percent correct scores, standard deviations, and ranges of scores on PEPS-C tasks for the three age groups. When performance for the three age groups was compared using a Kruskal Wallis ANOVA significant main effects of age on Chunking ($\chi^2(2, 45)=13.15, p=0.001$), Contrastive Stress ($\chi^2(2, 45)=13.14, p=0.001$), and PEPS-C total scores ($\chi^2(2, 45)=21.79, p=0.001$) were found. PEPS-C total scores were calculated as the average of the scores from the four prosody subtests. There were no effects of age group on Turn-end and Affect Reception scores (all $p>0.300$). Post-hoc Mann Whitney U tests (significance value set at $p<0.005 (0.05/9)$) showed that scores obtained by 7-8 year olds were significantly poorer than those obtained by 9-10 and 11-12 year olds for Chunking ($p \leq 0.003$), Contrastive Stress ($p \leq 0.003$), and PEPS-C total ($p=0.001$). There were no significant differences in scores obtained by the two older groups across PEPS-C tasks ($p \geq 0.072$). The PEPS-C data for the two older groups were therefore combined for further descriptive and statistical analyses. Mean percent correct scores obtained by 7-8 year old children on PEPS-C tasks were lower than the scores for the combined 9-12 year olds (Mage=10.99, SD=1.05, $n=31$; Figure 1). High standard deviations and wide ranges of scores obtained by the youngest group indicate greater intersubject variability in their performance (Figures 2 and 3). Compared to 7-8 year olds, smaller standard deviations and narrow ranges of scores were obtained by 9-12-year olds across the PEPS-C tasks. Most children (90%) in the 9-12 year old combined older age group performed above the chance level of 75%, with most achieving ceiling scores on the four PEPS-C subtests (Figure 3).

Outliers were present for three out of the four tasks for the older group, however. Thus, even though the majority of the children are successful at a task, there were five children (3 boys, 2 girls) performing very poorly compared to their peers. Ceiling effects were found for all tasks for some of the younger children. Among the 7-8 year olds, below chance level performance ($<75%$) occurred for one participant for the Turn-end and Affect Reception tasks and four participants for the Contrastive Stress Reception task.

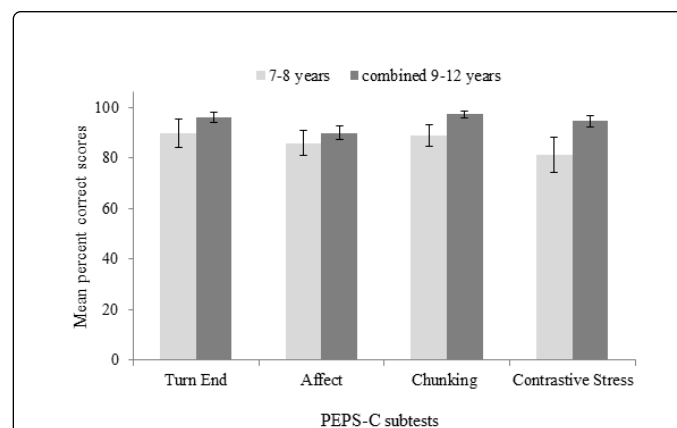


Figure 2: Means and 95% confidence intervals for PEPS-C subtests by age group.

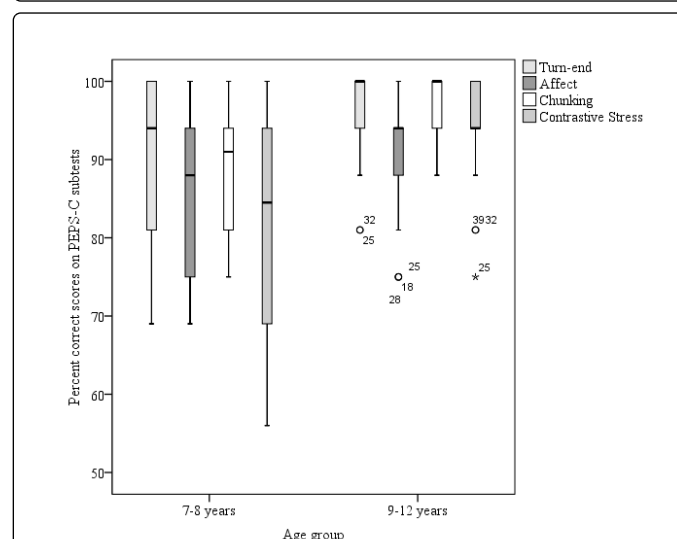


Figure 3: Box plots representing percent correct scores obtained by two age groups on PEPS-C subtests. The median scores are indicated by the thick horizontal line. Boxes indicate the data falling between the 25th and 75th percentile and the whiskers indicate the 95% confidence intervals.

Mann Whitney U tests were used to investigate differences in performance between the youngest (7-8 years) and the combined older age (9-12 years) group for the four PEPS-C subtests and PEPS-C total (significance value set at $p<0.01 (0.05/5)$). Scores obtained by 7-8 year olds were significantly poorer than those obtained by the combined 9-12 year olds for the Chunking ($U=80.00, p=0.001$), Contrastive Stress ($U=75.50, p=0.001$), and PEPS-C total ($U=27.50, p=0.001$; Table

2). There were no significant differences in scores obtained by the two groups for Turn-end ($U=140.50$, $p=0.043$) and Affect Reception tasks ($U=161.00$, $p=0.156$). These results match those obtained when the three age groups were compared.

Age	group	Turn-end	Affect	Chunking	Contrastive Stress	PEPS-C total
7-8 years (n=14)	M (SD)	89.85 (10.58)	85.92 (9.67)	88.92 (7.94)	81.42 (13.25)	86.53 (4.59)
	Mdn	94	88	91	84	87.75
	Range	69-100	69-100	75-100	56-100	78.25-92.25
9-10 years (n=16)	M (SD)	96.93 (5.10)	88.87 (8.35)	97.37 (3.77)	95.06 (6.48)	94.56 (4.00)
	Mdn	100	91	100	94	95.50
	Range	81-100	75-100	94-100	75-100	82.75-100
11-12 years (n=15)	M (SD)	95.13 (5.35)	91.00 (6.27)	97.20 (4.45)	94.26 (6.48)	94.40 (3.04)
	Mdn	94	94	100	94	94.00
	Range	88-100	81-100	88-100	81-100	89.00-98.50
Combined 9-12 years (n=31)	M (SD)	96.06 (5.22)	89.90 (7.38)	97.29 (4.05)	94.67 (6.38)	94.48 (3.51)
	Mdn	100	94	100	94	95.50
	Range	81-100	75-100	88-100	75-100	82.75-100

Table 2: Means, standard deviations, medians, and ranges of scores for PEPS-C subtests by age group.

Differences in performance based on PEPS-C prosodic task

Among the 7-8 year old children, there were no significant differences in scores across PEPS-C tasks ($\chi^2(3, 14)=5.347$, $p=0.148$). However, there were significant differences in scores among 9-12 year old children ($\chi^2(3, 31)=22.568$, $p=0.001$) depending on the task. Post-hoc analyses with Wilcoxon Signed-Rank tests (significance level set at $p<0.008$ (0.05/6)) showed that scores obtained by 9-12 year olds on the Affect Reception task were significantly poorer than those obtained on Turn-end ($Z=-3.106$, $p=0.002$) and Chunking Reception tasks ($Z=-3.856$, $p=0.001$). There were no significant differences in scores between any other pairs of tasks ($p \geq 0.012$).

Age group differences on DANVA 2 child paralinguistic subtest

Table 3 shows the percentage of errors made by three groups of children on two levels of emotion intensity and four different emotional categories. Overall more errors were made by 7-8 year-olds, followed by 9-10 year old children, with fewest errors made by 11-12 year olds. Kruskal Wallis ANOVAs were used to determine the effects of age on errors across the two levels of emotion intensity for the

DANVA 2 total error scores (four emotions combined, Tables 3 and 4). There were significant main effects of age for high emotion intensity errors ($\chi^2(2, 45)=6.831$, $p=0.033$), but not for low emotion intensity errors ($\chi^2(2, 45)=3.404$, $p>0.05$). Mann Whitney U tests (significance value set at $p<0.016$ (0.05/3)) showed that, for high emotion intensity, 7-8 year olds made more errors than 9-10 year olds ($U=51.00$, $p=0.008$) but did not differ from 11-12 year olds ($U=80.50$, $p>0.05$). Total scores for the two emotion intensities did not differ for the two older age groups and the performance of the younger age groups was the same as the older age groups for lower emotion intensity items ($p \geq 0.188$).

Age group	n	Happy		Sad		Angry		Fearful		Total	
		Low	High	Low	High	Low	High	Low	High	Low	High
7-8 years	14	29	14	40	21	19	5	33	19	30	15
9-10 years	16	27	4	27	6	25	2	31	13	28	6
11-12 years	15	22	6	22	16	16	4	29	18	22	14
Combined 9-12 years	31	25	5	25	11	20	3	30	15	50	17

Table 3: Percentage of errors for each age group across the four emotions and two emotion intensities (24 items in total, 12 per intensity, 6 per emotion) on DANVA 2 Child Paralinguistic subtest group.

Differences in DANVA 2 scores based on emotion intensity and emotional category

Wilcoxon Signed-Rank tests showed that total error scores for high emotion intensity items ($M=1.26$, $SD=1.23$) were significantly lower than the error scores for low emotion intensity items ($M=3.20$, $SD=1.60$, $Z=-4.984$, $p=0.001$; Table 5). Irrespective of the levels of emotion intensity, participants made more errors on items expressing fear, followed by sadness, then happiness, and had relatively few errors for anger (Table 3). Friedman ANOVA showed significant differences between emotional categories ($\chi^2(3, 45)=10.881$, $p=0.012$). Post-hoc analyses using Wilcoxon Signed-Rank tests revealed that the error scores obtained for fear stimuli were significantly higher than the error scores obtained for angry stimuli ($Z=-2.969$, $p=0.003$).

Wilcoxon Signed-Rank tests (significance value was set at $p<0.012$ (0.05/4)) were performed to determine the effects of emotion intensity on the errors obtained within the four emotion categories. There was no significant difference between high and low emotion intensity error scores for fear ($Z=2.439$, $p=0.015$; Table 4). Error scores for the other three emotion categories were lower for high emotion intensity (happiness: $Z=-3.774$, $p=0.001$; sadness: $Z=-2.641$, $p=0.008$; anger: $Z=-3.977$, $p=0.001$ (Table 5).

Age	Emotion Intensity	group	Happy	Sad	Angry	Fearful
7-8 years	Low	M (SD)	0.85 (0.86)	1.21 (1.05)	0.57 (0.75)	1.00 (1.03)

		Range	0-2		0-2	
	High	M (SD)	0.42 (0.64)	0.64 (0.84)	0.14 (0.36)	0.57 (0.75)
		Range	0-2		0-1	
9-12 years	Low	M (SD)	0.74 (0.68)	0.74 (0.81)	0.61 (0.66)	0.90 (0.83)
		Range	0-2		0-2	
	High	M (SD)	0.16 (0.37)	0.32 (0.59)	0.09 (0.30)	0.45 (0.56)
		Range	0-1		0-1	

Table 4: Means and standard deviations (error scores) for DANVA 2 Child Paralanguage subtest by emotion intensity (low and high) and emotion categories.

Emotion Intensity		Happy	Sad	Angry	Fearful	Total
Low	M (SD)	0.77 (0.73)	0.88 (0.91)	0.60 (0.68)	0.93 (0.88)	3.20 (1.60)
	Range	0-2	0-3	0-2	0-3	1-7
High	M (SD)	0.24 (0.48)	0.42 (0.69)	0.11 (0.31)	0.48 (0.62)	1.26 (1.23)
	Range	0-2	0-2	0-1	0-2	0-5

Table 5: Mean error scores and standard deviations on four emotional categories at two levels of emotion intensity for DANVA 2.

Emotion confusion matrix

Table 6 shows the emotion confusion matrix for the entire group of participants (N=45). The emotion that was most correctly identified was anger (88%), followed by happiness (83%), then sadness (78%), and finally fear (76%). Fear and sadness were the emotions that participants had the most difficulty identifying. Fear was most often confused with sadness (15% of the error responses for fearful tones were sad) and vice versa (12% of the error responses for sad tones were fearful). The confusion matrix shows that the errors were not randomly distributed, instead a clear pattern was observed where some pairs of emotions are confused with one another more often than others.

Stimulus	Response (%)			
	Happy	Sad	Angry	Fearful
Happy	82.96	8.88	3.33	4.81
Sad	6.29	78.14	3.33	12.22
Angry	1.48	8.88	88.14	1.48
Fearful	5.55	15.92	2.22	76.29

Note. The percentage of correctly identified emotions is given on the main diagonal in boldface type.

Table 6: Emotion confusion matrix for the entire group of participants (N=45) on DANVA 2 Child Paralanguage subtest (in proportion).

Gender differences

Mann Whitney U tests were performed to examine whether there were gender differences in performance on PEPS-C tasks and DANVA 2 subtest. No significant effects of gender were observed for any PEPS-C task (all $p > 0.868$; Table 7) or DANVA 2 subtest (all $p > 0.161$).

Gender	group	Turn-end	Affect	Chunking	Contrastive Stress	PEPS-C total
Girls (n=24)	M (SD)	94.37 (6.65)	90.00 (8.83)	93.04 (8.12)	90.75 (9.75)	92.04 (5.00)
	Mdn	94.00	94.00	94.00	94.00	92.37
	IQR	6.00	11.25	12.00	17.25	7.63
Boys (n=21)	M (SD)	93.85 (9.00)	87.14 (7.47)	96.57 (4.05)	90.33 (12.27)	91.97 (5.82)
	Mdn	100.00	88.00	100.00	94.00	94.00
	IQR	9.00	13.00	6.00	12.00	5.50

Note: IQR=Interquartile Range.

Table 7: Gender wise comparisons using PEPS-C scores.

Discussion

The PEPS-C results showed that 7-8 year olds performed significantly poorer than 9-12 year olds on Chunking and Contrastive Stress Reception tasks, indicating a developmental trend. The reduced standard deviation scores and narrow ranges of scores obtained by 9-12 year olds compared to the youngest group are also indicative of the age-related improvements. Moreover, most children in the oldest group achieved ceiling scores on the four PEPS-C subtests. Overall the results indicate that much of the age-related changes in prosody perception occur between 7 and 9 years. Previous studies using PEPS-C test have reported age-related improvements in receptive and expressive prosodic skills [23,52,55]. Wells et al. [23] reported significant developmental changes in prosodic abilities in children aged between 5 and 13 years. These results are consistent with Ludwig et al.'s (2014) findings that significant improvements in interaural and dichotic discrimination thresholds for acoustic parameters such as intensity, frequency, and signal duration occur between 6-7 and 8-9 years. Similarly, development effects on prosodic control have been reported based on acoustic analysis of prosody production and articulatory movement studies in children [1-3,34].

Even though a general age-related improvement in perception scores was observed across PEPS-C tasks, there were variations in the developmental pattern for different aspects of prosody. The older group performed significantly better than the 7-8 year olds on Chunking and Contrastive Stress Reception tasks. However, there were no significant differences between the older and younger age groups on Turn-end and Affect Reception tasks. This suggests that skills measured using PEPS-C Turn-end and Affect Reception subtests which involve discrimination of simple pitch movements are acquired in the early school-age period. While the PEPS-C Chunking subtest which requires judging speakers' use of timings cues and PEPS-C Contrastive Stress subtest which requires children to understand the use of accent/focus are acquired later and gradually. Previous studies have reported that comprehension of chunking and contrastive focus continues to develop up to 11 years [23,56]. Differential patterns in the

development of prosodic skills are supported by the prosody production literature for children. Grigos and Patel [34] investigated articulatory movements associated with the production of words with and without focus in 4, 7, and 11 year olds, and adults. Significant differences in duration, displacement, and velocity between focused and unfocused productions were seen between 7 and 11 year olds and adults, and there were differences between 11 year olds and adults. Grigos and Patel concluded that the ability to produce sentential stress starts to develop between seven and eleven years and continues throughout adolescence. Doherty, Fitzsimons, Asenbauer, and Staunton [57] examined prosody perception in typically developing children (N=40, aged between 5 and 9 years) using linguistic (discrimination of compound noun vs. noun phrase pairs and differentiation of questions/statements/commands) and affective prosody tasks. They found significant age-wise improvement in perceptual abilities up to 8; 5 years. They also reported that vocal emotion recognition in children develops later than the corresponding linguistic ability. Ito, Bibyk, Wagner, and Speer [58] reported age-related improvements in interpreting contrastive accent in children aged between 6 and 11 years, however even the 11 year olds showed delayed responses compared to adults. This suggests that it may take many years for children to acquire the pragmatic meaning of pitch accent. Early mastery of question-statement distinction over contrastive stress patterns could be related to greater exposure and familiarity effects. The infant directed speech literature suggests that motherese includes large amount of emotional information and utterances in the form of question-statement [59-61]. In conversational English, contrastive stress usually occurs in the final word position of a sentence while the PEPS-C Contrastive Stress task uses stress on different word positions (e.g., I wanted a BLUE and green socks (emphasis on the first colour) vs. I wanted a blue and GREEN socks (emphasis on the second colour)). This may not be the familiar pattern for children and hence greater access to auditory cues may be crucial to make this distinction. This is further corroborated by Balogh, Swinney, and Tigue's [62] findings that the ability to respond to contrastive stress is related to a general sensitivity to prosodic cues and is distinct from syntactic and pragmatic knowledge.

There were no significant differences in performance within the 7-8 year olds across the PEPS-C tasks, however performance on PEPS-C Affect Reception task was significantly poorer than that for Turn-end and Chunking Reception tasks for the 9-12 year olds. This suggests that the PEPS-C Affect Reception task was the most difficult for the 9-12 year olds compared to other PEPS-C tasks. This could be because the PEPS-C Affect Reception task uses a single word test items (names of food items) rather than a sentence context which is less likely to happen in real life situations (less ecological validity; Diehl & Paul [5]). The DANVA 2 Child Paralanguage subtest results provide a comprehensive view of affective prosody perception abilities in children. DANVA 2 uses sentence level stimuli to assess perception of four different emotions (happy, sad, angry, and fearful) whereas the PEPS-C Affect Reception subtest includes only two emotions (like/dislike). There were no gender effects on PEPS-C or DANVA 2 subtest performance. This is consistent with the results reported by Wells et al. [23] and Peppé et al. [63].

DANVA 2 Child Paralanguage subtest results showed that 7-8 year olds made more errors, followed by 9-10 year olds, and least number of errors was made by 11-12 year olds. These results suggest a developmental trend in affective prosody perception abilities in children using DANVA 2 subtest; however this did not reach statistical significance. Nowicki and Duke [7] reported significant age-related

changes in 6-10 year olds on DANVA 2 Child Paralanguage subtest. They also reported a strong correlation between vocal emotion recognition and academic achievement in children while DANVA 2 facial expression and posture recognition subtests did not show any correlation. Significant correlations between vocal emotion recognition and social adjustment (measured using Social Dysfunction Index) in adults with schizophrenia were reported by Hooker and Park [64]. Unfortunately, emotion processing in children has been mainly assessed through visual modality by using facial expression tasks, and not much focus has been given to vocal emotion recognition. This is of concern because the auditory system matures earlier than the visual system [64,65] and understanding of vocal emotion expressions plays a major role in early emotional development [38,67]. Halberstadt & Eaton [68] reported that reduced family expressiveness of emotions through facial expressions and voice were associated with poor emotion understanding and expression in children. Early aberrations in emotion processing need to be identified and treated in order to ensure normal social and emotional development.

Overall the DANVA 2 results indicate that the errors obtained for different emotions varied considerably depending on the level of emotional intensity. Emotions presented at high intensities were recognised significantly better than those presented at low intensities for all emotions, except for fear. These findings are consistent with the results of Juslin and Laukka [47] who reported that listeners were able to decode happiness, sadness, anger, fear, and disgust vocal emotions presented at strong emotion intensity better than for weak emotion intensity. This is further supported by Bänziger and Scherer's [69] findings that there is an increase in F0 mean and F0 range with increasing intensity, which serves as a cue for easier detection of high emotion intensity stimuli. They reported that F0 parameters like mean, range, and minimum and maximum F0 peak for low emotion intensities - such as 'sadness', 'calm joy', and 'anxious fear' are generally lower than the F0 values for emotions with high intensities such as 'despaired sadness', 'elated joy', 'panic fear', and 'hot anger'. It is important to know how well children understand low emotion intensity cues, as in real life situations expressions of emotions are often subtle [46]. Emotion intensity has not been systematically varied in studies comparing atypical and typical populations. This is an important issue because emotion processing difficulties in atypical populations may be underestimated if only high intensity stimuli are used. Considering the level of emotion intensity as a factor is useful in identifying typical error patterns associated with different disorders [40,70,71]. Baum and Nowicki [45] reported that accurate perception of low emotion intensity cues, but not high intensity cues, was related to social competence in typically developing children. These findings indicate the importance of assessing prosody perception at different intensity levels in typically developing children in order to have a basis for evaluating children with disordered prosody.

The lowest accuracy was observed for fearful emotions followed by sadness. Highest accuracy was noted for angry followed by happiness, consistent with the results from previous studies [27,47]. Bänziger and Scherer [69] reported specific differences in F0 contours for different emotion categories that make certain emotions easier to identify than others. For emotions such as 'hot anger', 'cold anger', and 'elation joy' the F0 excursions in the second part of the utterance tend to be larger than for sadness or happiness. The shape of the F0 contour also changes depending on the emotion category; steeper final falls were observed for anger compared to a progressive decrease (sadness) and increase (happiness) in F0 until the final fall. The additional F0 information associated with anger and happiness could be the reasons

why these emotions are perceived more accurately than others by the children. Most of the confusions between emotions reported in the present study can be described as symmetrical (a term borrowed from Juslin & Laukka [47]). For example, sadness was often confused with fear, and fear was confused with sadness. The same is true for sad and happy emotions and fear and happy emotions. These confusions mostly occurred for low emotion intensity items; suggesting that subtle acoustic cues are insufficient to accurately discriminate different emotions [47,48,69]. Asymmetrical confusions were also present, such as anger was mostly confused with sadness, but sadness was rarely confused with anger. However, sadness was the most frequently chosen incorrect alternative. There is minimal research to suggest that there are developmental differences in understanding vocal emotions depending on the emotion categories [73,74]. Further research should investigate the mechanisms by which children develop abilities to recognize different emotions.

Conclusion

The present study revealed a number of significant findings regarding prosody perception abilities in typically developing 7-12 year old children. Four receptive prosody subtests of PEPS-C and Child Paralanguage subtest of DANVA 2 were used. This research provided normative data for PEPS-C receptive prosody subtests and reported that development of receptive prosodic skills occurs between seven and nine years. A differential pattern of development for different aspects of prosody was found; chunking and contrastive stress reception skills develop at a later age compared to turn-end and affect recognition. Age-related improvements in performance on DANVA 2 subtest were observed; however these did not reach statistical significance. DANVA 2 scores varied depending on the level of emotion intensity, with high emotion intensity stimuli perceived more accurately than low emotion intensity items and this was consistent across the emotions, except for fear. There were no gender effects on PEPS-C or DANVA 2 scores. The results have clinical implications for assessing prosody perception abilities in atypical populations.

References

1. Grigos M, Patel R (2007) Articulator movement associated with the development of prosodic control in children. *J Speech Lang Hear Res* 50: 119-130.
2. Patel R, Brayton JT (2009) Identifying prosodic contrasts in utterances produced by 4-, 7-, and 11-year-old children. *J Speech Language Hear Res* 52: 790-801.
3. Patel R, Grigos MI (2006) Acoustic characterization of the question-statement contrast in 4, 7, and 11 year-old children. *Speech Commun* 48: 1308-1318.
4. Snow D (1994) Phrase-final lengthening and intonation in early child speech. *J Speech Hear Res* 37: 831-840.
5. Diehl JJ, Paul R (2009) The assessment and treatment of prosodic disorders and neurological theories of prosody. *Int J Speech Lang Pathol* 11: 287-292.
6. Peppé S, McCann J (2003) Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clin Linguist Phon* 17: 345-354.
7. Nowicki S Jr, Duke MP (1994) Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. Special Issue: Development of nonverbal behaviour: II Social development and nonverbal behaviour. *J Nonverbal Behav* 18: 9-35.
8. Mehler J, Jusczyk P, Lambertz G, Halsted N, Bertoincini J et al. (1988) A precursor of language acquisition in young infants. *Cognition* 29: 143-178.
9. Nazzi T, Jusczyk PW, Johnson EK (2000) Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *J Memory Lang* 43: 1-19.
10. Jusczyk PW (1993) From general to language-specific capacities: The WRAPSA model of how speech perception develops. *J Phon* 21: 3-28.
11. Gilbert H, Robb M (1996) Vocal fundamental frequency characteristics of infant hunger cries: Birth to 12 months. *Int J Pediatr Otorhinolaryngol* 34: 237-243.
12. Lind K, Wermke K (2002) Development of the vocal fundamental frequency of spontaneous cries during the first 3 months. *Int J Pediatr Otorhinolaryngol* 64: 97-104.
13. Gerken LA, Jusczyk PW, Mandel DR (1994) When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition* 51: 237-265.
14. Hirsh-Pasek K, Kemler Nelson D, Jusczyk PW, Wright Cassidy K, Druss B (1987) Clauses are perceptual units for prelinguistic infants. *Cognition* 26: 269-286.
15. Jusczyk PW, Hirsch-Pasek K, Kemler Nelson DG, Kennedy LJ, Woodward A (1992) Perception of acoustic correlates of major phrasal units by young infants. *Cogn Psychol* 24: 252-293.
16. Kemler Nelson D, Hirsh-Pasek K, Jusczyk PW, Wright Cassidy K (1989) How prosodic cues in motherese might assist language learning. *J Child Lang* 16: 53-68.
17. Nazzi T, Kemler Nelson DG, Jusczyk PW, Jusczyk AM (2000) Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy* 1: 123-147.
18. Turk A, Jusczyk PW, Gerken LA (1995) Do English-Learning Infants use Syllable Weight to Determine Stress? *Lang Speech* 38: 143-158.
19. Jusczyk PW, Cutler A, Redanz N (1993) Infants' sensitivity to predominant word stress patterns in English. *Child Dev* 64: 675-687.
20. Cutler A, Swinney D (1987) Prosody and the development of comprehension. *J Child Lang* 14: 145-167.
21. Snow D (1998) Prosodic markers of syntactic boundaries in the speech of 4-year old children with normal and disordered language and development. *J Speech Lang Hear Res* 41: 1158-1170.
22. Roach P (2000) English phonetics and phonology- A practical course. Cambridge University Press, United Kingdom.
23. Wells B, Peppé S, Goulandris N (2004) Intonation development from five to thirteen. *J Child Lang* 31: 749-778.
24. Christophe A, Gout A, Peperkamp S, Morgan J (2003) Discovering words in the continuous speech stream: The role of prosody. *J Phon* 31: 585-598.
25. Johnson EK, Jusczyk PW (2001) Word segmentation by 8-month-olds: When speech cues count more than statistics. *J Memory Lang* 44: 548-567.
26. Port RF (2003) Meter and speech. *J Phon* 31: 599-611.
27. Johnstone T, Scherer KR (2000) Vocal communication of emotion. In: Lewis M, Harviland-Jones JM (Edn.), *Handbook of emotions*. New York: Guilford Press, pp. 220-235.
28. Murray IR, Arnott JL (1993) Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *J Acoust Soc Am* 93: 1097-1108.
29. d'Alessandro C (2006) Voice source parameters and prosodic analysis. In: Sudhoff S et al., (Edn.), *Methods in Empirical Prosody Research*, Walter De Gruyter, Berlin, New York, pp. 63-87.
30. Couper-Kuhlen E, Selting M (1996) *Prosody in conversation: Interactional studies*. Cambridge University Press, Cambridge.
31. Schegloff EA (1996) Turn organization: One intersection of grammar and interaction. In: Ochs E, Schegloff EA, Thompson SA (Edn.) *Interaction and grammar*. Cambridge University Press, Cambridge, pp. 52-133.
32. De Ruiter LE (2014) How German children use intonation to signal information status in narrative discourse. *J Child Lang* 41: 1015-1061.
33. Wonnacott E, Watson DG (2008) Acoustic emphasis in 4-year-olds. *Cognition* 107: 1093-1101.

34. Grigos M, Patel R (2010) Acquisition of articulatory control for sentential focus in children. *J Phon* 38: 706-715.
35. Potamianos A, Narayanan S (1998) Spoken dialog systems for children. Proceedings of the Int Conference on Acoustics Speech Signal Processing, Washington 1: 197-200. doi: 10.1109/ICASSP.1998.674401.
36. Burnham D (1993) Visual recognition of mother by young infants: Facilitation of speech. *Perception* 22: 1133-1153.
37. Fernald A (1993) Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Dev* 64: 657-674.
38. Dunn J (2003) Emotional development in early childhood: A social relationship perspective. In: Davidson RJ et al., (Edn.), *Handbook of Affective Sciences*. Oxford Press, New York, pp. 332-346.
39. Rosnay MD, Fink E, Begeer S, Slaughter V, Peterson C (2014) Talking theory of mind talk: young school-aged children's everyday conversation and understanding of mind and emotion. *J Child Lang* 41: 1179-1193.
40. Mazefsky CA, Oswald DP (2007) Emotion perception in Asperger's syndrome and high-functioning autism: The importance of diagnostic criteria and cue intensity. *J Autism Dev Disord* 37: 1086-1095.
41. Mazefsky CA (2002) Emotion perception in Asperger's Syndrome and High-Functioning Autism: The importance of diagnostic criteria and cue intensity (Doctoral dissertation, Virginia Commonwealth University).
42. Achenbach TM (1991) Integrative guide for the 1991 CBCL/4-18, YSR and TRF profiles. Burlington: Department of Psychiatry, University of Vermont.
43. Achenbach TM, Rescorla LA (2001) The manual for the ASEBA school-age forms & profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
44. Bruininks RH, Woodcock RW, Weatherman RE, Hill BK (1996) Scales of Independent Behavior-Revised. Riverside Publishing, Itasca IL.
45. Baum KM, Nowicki S (1998) Perception of emotion: Measuring decoding accuracy of adult prosodic cues varying intensity. *J Nonverbal Behav* 89: 89-107.
46. Russell JA, Barrett LF (1999) Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *J Pers Soc Psychol* 76: 805-819.
47. Juslin PN, Laukka P (2001) Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion* 1: 381-412.
48. Bance R, Scherer KR (1996) Acoustic profiles in vocal emotion expression. *J Personality Social Psychol* 70: 614-636.
49. Juslin PN, Laukka P (2003) Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin* 129: 770-814.
50. Luo X, Fu Q-J, Galvin JJ (2007) Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends Amplif* 11: 301-315.
51. Wells B, Peppé S (2003) Intonation abilities of children with speech and language impairments. *J Speech Language Hear Res* 46: 5-20.
52. Gibbon FE, Smyth H (2013) Preschool children's performance on Profiling Elements of Prosody in Speech-Communication (PEPS-C). *Clin Linguist Phon* 27: 428-434.
53. Peppé S (2009) Aspects of identifying prosodic impairment. *Int J Speech-Lang Pathol* 11: 332-338.
54. Howell DC (2014) *Fundamental statistics for the behavioural sciences*. Cengage Learning, Wadsworth.
55. Foley M, Gibbon F, Peppé S (2011) Benchmarking typically developing children's prosodic performance on the Irish-English version of the Profiling Elements of Prosody in Speech-Communication (PEPS-C). *J Clin Speech Lang Stud* 18: 19-40.
56. Cruttenden A (1985) Intonation comprehension in ten-year-olds. *J Child Lang* 12: 643-661.
57. Doherty CP, Asenbauer FB, Staunton H (1999) Discrimination of prosody and music by normal children. *Eur J Neurol* 6: 221-226.
58. Ito K, Bibyk SA, Wagner L, Speer SR (2014) Interpretation of contrastive pitch accent in six- to eleven-year-old English-speaking children (and adults). *J Child Lang* 41: 84-110.
59. Durkin K, Rutter DR, Tucker H (1982) Social interaction and language acquisition: Motherese help you. *First Lang* 3: 107-120.
60. Soderstrom M, Blossom M, Foygel R, Morgan JL (2008) Acoustical cues and grammatical units in speech to two preverbal infants. *J Child Lang* 35: 869-902.
61. Trainor LJ, Austin CM, Desjardins RN (2000) Is infant-directed speech prosody a result of the vocal expression of emotion? *Psycholog Sci* 11: 188-195.
62. Balogh JE, Swinney D, Tigue Z (1998) Sensitivity to contrastive stress: A study of individual differences. Paper presented at the 11th CUNY Conference on Human Sentence Processing, New Brunswick, Quebec.
63. Peppé S, Maxim J, Wells B (2000) Prosodic variation in Southern British English. *Lang Speech* 43: 309-334.
64. Hooker C, Park S (2002) Emotion processing and its relationship to social functioning in schizophrenic patients. *Psychiatry Res* 112: 41-50.
65. Gottlieb G (1971) *Development of species identification in birds: An inquiry into the prenatal determinants of perception*. University of Chicago Press, Chicago.
66. Jusczyk PW (1998) *The discovery of spoken language*. MIT Press, Cambridge, MA.
67. Denham SA (1998) *Emotional development in young children*. Guilford Press, New York.
68. Halberstadt AG, Eaton KL (2002) A meta-analysis of family expressiveness and children's emotion expressiveness and understanding. *Marriage Fam Rev* 34: 35-62.
69. Bänziger T, Scherer KR (2005) The role of intonation in emotional expressions. *Speech Commun* 46: 252-267.
70. Castelli F (2005) Understanding emotions from standardized facial expressions in autism and normal development. *Autism* 9: 428-449.
71. Grossman RB, Tager-Flusberg H (2012) "Who said that?" Matching of low- and high-intensity emotional prosody to facial expressions by adolescents with ASD. *J Autism Dev Disord* 42: 2546-2557.
72. Cooper WE, Eady SJ, Mueller PR (1985) Acoustical aspects of contrastive stress in question-answer contexts. *J Acoust Soc Am* 77: 2142-2156.
73. Mandel DR, Jusczyk PW, Kemler Nelson D (1994) Does sentential prosody help infants organise and remember speech information? *Cognition* 53: 155-180.