Child Development, May/June 2001, Volume 72, Number 3, Pages 834–843

Children's Understanding of Emotion in Speech

J. Bruce Morton and Sandra E. Trehub

Children's understanding of emotion in speech was explored in three experiments. In Experiment 1, 4- to 10-year-old children and adults ($N = 165$) judged the happiness or sadness of the speaker from cues conveyed by propositional content and affective paralanguage. When the cues conflicted (i.e., a happy situation was described with sad paralanguage), children relied primarily on content, in contrast to adults, who relied on paralanguage. There were gradual developmental changes from 4-year-olds' almost exclusive focus on content to adults' exclusive focus on paralanguage. Children of all ages exhibited greater response latencies to utterances with conflicting cues than to those with nonconflicting cues, indicating that they processed both sources of emotional information. Children accurately labeled the affective paralanguage when the propositional cues to emotion were obscured by a foreign language (Experiment 2, $N = 20$) or by low-pass filtering (Experiment 3, $N = 60$). The findings are consistent with children's limited understanding of the communicative functions of affective paralanguage.

INTRODUCTION

Information about a speaker's feelings is carried in two vocal channels: propositional content and paralanguage. Speakers can use propositional content, or the constituent words and phrases of an utterance, to depict their feelings directly (e.g., "I feel happy") or to describe a situation that has positive or negative emotional implications (e.g., "I won a trip to Paris"). They can also express their emotions paralinguistically by altering their speaking rate, pitch level, pitch contours, and voice quality (Frick, 1985). If the content is emotionally neutral (e.g., "Today is Wednesday"), a speaker's feelings (neutral, positive, or negative) can be discerned from the paralanguage. For example, speech marked by high pitch, rapid tempo, large pitch range, and bright voice quality signals happiness; speech with low pitch, slow tempo, narrow pitch range, and soft voice signals sadness (Scherer, 1986). Although some paralinguistic features are used intentionally to convey attitude (i.e., the speaker's appraisal of the situation described; Bolinger, 1986; Levy, 1984), others are by-products of the speaker's emotion or arousal (Bachorowski & Owren, 1995; Scherer, 1986). Whereas paralinguistic cues to attitude vary across cultures, those linked to basic emotions are universal (Frick, 1985).

Typically, paralinguistic cues reinforce emotive verbal content, as when a happy event is described in a joyful manner. At times, however, propositional and paralinguistic cues convey contradictory cues to emotion. How might listeners judge a speaker's feelings in the context of discrepant propositional and paralinguistic information? The available evidence indicates that adults consider all available cues (Reilly & Muzekari, 1979), but that they rely primarily on paralanguage in such circumstances (Mehrabian & Wiener, 1967). Children's relative weighting of propositional and paralinguistic cues to emotion is likely to differ at different ages.

Paralinguistic cues to emotion seem to be salient and meaningful from the earliest months of life. Well before propositional cues are interpretable, the affectively charged paralanguage of infant-directed speech captures infants' attention more effectively than does the neutral paralanguage of typical adult-directed speech (e.g., Cooper & Aslin, 1990; Fernald, 1985; Werker & McLeod, 1989). Rather than responding to infant-directed speech per se, infants may respond on the basis of its heightened emotionality. Indeed, infant- and adult-directed messages with comparable emotional intent (e.g., love/comfort) are surprisingly similar in their prosodic features (Trainor, Austin, & Desjardins, 2000). Infants not only respond to such vocal cues to emotion; they also respond in affectively appropriate ways. For example, 4- and 5-month-old infants react more positively to messages with positive emotional valence than to those with negative emotional valence (Fernald, 1993; Papoušek, Bornstein, Nuzzo, Papoušek, & Symmes, 1990). Findings such as these lend credence to the notion that some prosodic aspects of speech are intrinsically meaningful (Fernald, 1992).

Although prelinguistic infants are capable of differentiating various vowels and consonants (for a review, see Jusczyk, 1997), prosodic cues command

© 2001 by the Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2001/7203-0015

much of their attention in everyday life (Fernald, 1991; Lewis, 1951). Once language acquisition is underway, however, new sources of information about affect emerge. For example, toddlers acquire emotion words like "happy" and "sad" (Bretherton & Beeghly, 1982; Bretherton, Fritz, Zahn-Waxler, & Ridgeway, 1986; Ridgeway, Waters, & Kuczaj, 1985). By their fourth year, children engage in simple dialogues about the causes and consequences of emotions (Bretherton & Beeghly, 1982; Dunn, Brown, & Beardsall, 1991). When given descriptions of familiar situations, they infer appropriate emotional consequences (Borke, 1971).

The achievement of the aforementioned milestones depends on young children's ability to deploy their attention in selectively appropriate ways. Nevertheless, limited attentional resources seem to constrain children's coordination of multiple cues in social reasoning situations (Burns & Cavey, 1957; Greenspan, Barenboim, & Chandler, 1976; but see Gnepp, 1983). Such limitations also influence the course of expressive and receptive language development. For example, 13-month-olds are unable to coordinate word production with the expression of vocal affect (Bloom, 1993; Bloom & Beckwith, 1989). Although they are capable of emotive vocalizations, their vocal affect is predominantly neutral when producing words. Moreover, 3- and 4-year-olds' stories include rich affective prosody but impoverished content, in contrast to 7- and 8-year-olds, whose coherent narratives are affectively neutral (Bamberg & Reilly, 1996). By 10 years of age, children's expressive prosody complements their story content.

Attentional constraints are also evident in children's interpretation of spoken language. For example, first graders have difficulty accessing the dual meanings of simple homonyms (e.g., *bark*), remaining fixated on the first meaning considered (Cramer, 1983; but see Backscheider & Gelman, 1995). Even 9-year-olds have difficulty ignoring salient content cues ("The cats are barking so loudly.") when making judgments about syntactic form (Bialystok, 1993).

How might these factors affect judgments of speaker affect from messages with discrepant propositional and paralinguistic cues? On the one hand, it is clear that even prelinguistic listeners are sensitive to paralanguage. On the other hand, age-related constraints on attention may lead children to accord greater weight to the more salient dimension. In their daily interactions with parents and peers, propositional cues are more informative, or salient, than paralinguistic cues. Although content is also salient for adults, they would likely remain vigilant for paralinguistic cues that shed light on the speaker's attitude, mood, and believability.

In fact, there is evidence that children accord greater weight to content than to paralanguage when judging the attitude of a speaker (Bugental, Kaswan, & Love, 1970; Friend & Bryant, 2000; Solomon & Ali, 1972). Friend and Bryant (2000) presented utterances with positive or negative content combined with happy or angry paralanguage to 4-, 7-, and 10-year-olds, who judged whether the speaker felt happy or "mad." The responses of 4- and 7-year-olds reflected greater weighting of content than of paralanguage, whereas those of 10-year-olds reflected greater weighting of paralanguage. The use of happiness and anger as the target emotions is somewhat problematic because of common acoustic cues such as high pitch, large pitch excursions, rapid tempo, and loud voice (Scherer, 1986). Indeed, children sometimes confuse happy and angry paralanguage; interestingly, they do not confuse angry and sad paralanguage, despite their common negative valence (Dimitrovsky, 1964). Moreover, the test utterances expressed favorable or unfavorable opinions about an addressee (e.g., "You'll never behave yourself"; "You play very well") rather than describing situations that had clear emotional consequences for the speaker. Thus, it is unclear whether children would show comparable reliance on propositional cues in messages that provided unambiguous propositional and paralinguistic cues to the speaker's feelings.

The principal goal of the present investigation was to document age-related changes in the understanding of speaker affect in the context of messages with conflicting propositional and paralinguistic cues. Participants listened to sentences that described happy or sad events involving the speaker (e.g., "My mommy gave me a treat"; "My dog ran away from home") but that excluded specific emotion words (e.g., happy, sad). Each sentence was presented once with happy paralanguage and once with sad paralanguage. Happiness and sadness were selected as the target emotions because of their familiarity to young children and their distinct acoustic cues (Scherer, 1986). Participants were instructed to listen to the speaker's voice and to judge whether she felt happy or sad.

If children consistently responded on the basis of one dimension, either propositional content or paralanguage, it would be difficult to confirm that they had processed the other dimension. Their processing of both dimensions would be revealed, however, by greater response latency to messages with conflicting cues compared with those with nonconflicting cues. Reaction time measures have revealed young children's detection of other types of conflicting cues (Beal & Flavell, 1982; Bearison & Levey, 1977; Jerger, Martin, & Pirozzolo, 1988). Children's explicit aware-

ness of the conflicting cues was probed by means of a series of questions. Comparable interview techniques have been used to assess children's emotional reasoning (Deutsch, 1974; Gnepp, 1983; Greenspan et al., 1976; Hoffner & Badzinski, 1989) and their understanding of ambiguity (Beal & Flavell, 1982).

On the basis of previous research (Bugental et al., 1970; Friend & Bryant, 2000; Mehrabian & Wiener, 1967; Solomon & Ali, 1972; Solomon & Yaeger, 1969), children were expected to give greater weight to propositional than to paralinguistic cues when the cues conflicted. Adults were expected to do the opposite. Thus, children were expected to judge a speaker's feelings by what she said and adults by how she spoke.

EXPERIMENT 1

Method

Participants. There were 165 participants: twenty-one 4-year-olds (*range* = 4,0–4,3, *M* = 4,1), twenty 5-year-olds (*range* = 5,0–5,3, *M* = 5,1), twenty-one 6-year-olds (*range* = 6,0–6,3, *M* = 6,1), twenty-two 7-year-olds (*range* = 7,0–7,3, *M* = 7,1), twenty-one 8-year-olds (*range* = 8,0–8,3, *M* = 8,2), twenty 9-year-olds (*range* = 9,0–9,3, *M* = 9,2), twenty 10-year-olds (*range* = 10, 0–10,3, *M* = 10,2), and twenty young adults (*range* = 18–22, *M* = 19,5). The children were drawn from families in the local community who had volunteered to participate in campus research. The adults were students in an introductory psychology course who received partial course credit for their participation. All participants spoke English fluently and were from middle-class backgrounds. Two additional participants were tested but were excluded because they made more than four errors on trials with consistent cues (e.g., responding "sad" to an utterance with happy content and paralanguage).

Apparatus and stimuli. Testing took place in a quiet room, 3 × 4 m in size, with a Power Macintosh 8100 computer. Children sat facing the monitor, and the experimenter sat beside the child. The experimenter used the computer keyboard to initiate trials and record the child's responses. Adults initiated trials and recorded their own responses by means of the computer keyboard.

The stimuli consisted of 40 spoken utterances. Sentences describing happy (*n* = 10) and sad (*n* = 10) situations (see Table 1) were each recorded twice, once with happy paralanguage (higher pitch level, greater pitch and loudness variation, and a faster speaking rate as compared with sad paralanguage), and once with sad paralanguage (lower average pitch, attenuated pitch and loudness variation, and slower speaking rate relative to happy paralanguage).[1] The result was 20 utterances in which propositional and paralinguistic cues to emotion were consistent and 20 sentences in which these cues conflicted. An additional four utterances with consistent propositional and paralinguistic cues were used as practice stimuli. All utterances were spoken by the same woman and were digitally recorded by means of a Radius computer and SoundScope software (GW Instruments, Inc.). Utterances were presented at a comfortable listening level, as determined by pilot testing. The order of the test materials was randomized for each participant. To ensure that potential differences in response time to conflicting and nonconflicting utterances were not attributable to differences in utterance duration, the duration of all test sentences was compared by means of Sound Designer software (GW Instruments, Inc.). Duration did not differ for sentences with conflicting (1.99 s, *SD* = 0.73) and nonconflicting (2.15 s, *SD* = 0.89) cues, $t(38) = 0.57$, $p = .57$.

Table 1 Happy and Sad Sentences

Happy sentences
1. My mommy gave me a treat.
2. My soccer team just won the championship.
3. I got an ice cream for being good.
4. I came in first place in a race today.
5. Dad gave me a new bike for my birthday.
6. I am having a party and all my best friends are coming.
7. My teacher says that I'm the smartest in the class.
8. I had my favorite cake for dessert.
9. Grandmother told me I'm very special.
10. I won a prize for being the fastest swimmer.
Sad sentences
1. My dog ran away from home.
2. My bike is broken so I can't go riding with my friends.
3. I lost my baseball glove today.
4. I lost my sticker collection.
5. I am not allowed to go outside and play with my friends.
6. My best friend doesn't like me anymore.
7. I fell off my bike and everyone made fun of me.
8. I lost the toy that Grandmother gave me for Christmas.
9. All the kids at camp tease me.
10. I lost all my money on the way to the store.

Procedure. Participants were tested individually. Children were told that they were going to be playing a short listening game. The experimenter's friend, Marianne, would be talking about many different things. Children were instructed to listen carefully to Marianne's voice and to say whether she was feeling happy or sad. After children successfully completed

[1] Sample auditory stimuli are available on the Web (http://psych.utoronto.ca/~bruce/Sounds/html).

the four practice trials (i.e., judging "happy" utterances as happy and "sad" utterances as sad) without feedback, the experimenter repeated the instructions and proceeded to the test phase. Response latency, defined as time from onset of the test sentence to entry of the response ("happy" or "sad"), was recorded automatically. No feedback was provided except for noncontingent encouragement. After responding to all sentences, children were asked if there was anything silly or weird about the way Marianne sometimes talked. Those who responded "yes" were asked to describe what they found silly. They were then asked whether there were any "tricks" in the listening game; those who responded "yes" were asked to describe the tricks. They were also asked whether Marianne had done a good job expressing her feelings. Those who replied "no" were asked to elaborate. Finally, they were asked whether it was easy to figure out how Marianne was feeling. Those who replied "no" were asked for details. These questions were inadvertently excluded for three 4-year-olds and six 5-year-olds.

Adult participants received written instructions. They were told that they would be hearing 40 short sentences. For each sentence, they were to listen carefully to the speaker's voice, indicating whether she was feeling happy or sad by pressing "h" or "s" on the keyboard. Adults were instructed to rest their left index finger on the "s" key and their right index finger on the "h" key so that they could respond quickly and accurately. After four training trials, adults read a summary version of the instructions, after which they proceeded to the test phase. At the conclusion of the task, adults were asked whether there was anything unusual about some of the utterances. Those who responded "yes" were required to indicate what they found unusual.

Results

Judgments of the 20 test utterances with conflicting cues were assigned scores as follows. Interpretations based on content received a score of 1; interpretations based on paralanguage received a score of 0. Thus, participants who attended exclusively to content received a score of 20; those who attended exclusively to paralanguage received a score of 0. Because scores in each age group were not normally distributed, they were grouped into one of three categories: Paralinguistic Focus (0–6), Mixed Focus (7–13), and Content focus (14–20). The resultant age by category frequency distribution is shown in Table 2. It is apparent that adults responded exclusively on the basis of paralanguage and that 4-year-olds responded primarily to content. Between 5 and 10 years of age, children showed gradual diminution in their reliance on content. Across all age groups, few participants scored in the intermediate range (Mixed Focus), indicating that they focused either on content or on paralanguage. A chi-square analysis confirmed the significant interaction between age and category, $\chi^2(14, N = 165) = 72.5$, $p < .01$. This interaction remained significant when boys and girls were considered separately, and when adult scores were removed, $\chi^2(12, N = 145) = 29.42$, $p < .05$.

Table 2 Children's and Adults' Responses to Sentences with Conflicting Cues

Age (years)	Paralinguistic Focus (scores of 0–6)	Mixed Focus (scores of 7–13)	Content Focus (scores of 14–20)
4	1	3	17
5	2	1	17
6	3	4	14
7	4	1	17
8	7	0	14
9	7	4	9
10	11	1	8
Adults	20	0	0
Totals	55	14	96

Differential response latency was calculated as the difference in total response time between the first four sentences with conflicting cues and the first four sentences with nonconflicting cues. (Hesitations to sentences with conflicting cues—reflecting surprise or uncertainty—were expected to decrease with increasing exposure to such sentences.) Differential response latency, converted to percentage change (to permit comparisons across age groups), significantly exceeded chance levels (0) for all groups (see Table 3), as revealed by t tests. (As expected, this difference decreased over time for all groups.) A one-way ANOVA

Table 3 Latency Differences between Sentences with Conflicting and Nonconflicting Cues

Age (years)	Mean Difference (ms)	Percentage Change
4	1904.6	31
5	878.0	18
6	876.9	21
7	2614.5	57
8	1430.8	34
9	1721.9	42
10	1188.2	31
Adults	363.5	13

Table 4 Proportions of Participants Describing the Ambiguous Cues and Commending the Speaker's Communication Efficacy

Age (years)	Description of Ambiguity	Favorable Evaluation
4	.00	.83
5	.08	1.00
6	.65	.85
7	.77	.63
8	.95	.53
9	1.00	.15
10	1.00	.15
Adults	1.00	N.A.

Note: N.A. = not applicable.

revealed no latency differences (i.e., percentage change) between age groups, $F(7, 156) = 2.05$, *ns*.

Responses to follow-up questions were coded for participants' ability to describe the unusual aspect of some test sentences. Participants who alluded to the speaker saying happy things but sounding sad (or saying sad things but sounding happy) were coded as describing the conflict. Those who acknowledged that some sentences seemed weird or silly without providing appropriate justification were coded as not describing the conflict. Responses as a function of age are shown in Table 4. (As noted, these questions were not posed to three of the 4-year-olds and six of the 5-year-olds.) Table 4 also shows the proportion of children who claimed that the speaker had expressed her feelings well. As can be seen in Table 4, most 4- and 5-year-olds failed to describe the conflicting cues. In fact, these children typically denied that the speaker ever talked in a silly way. By 7 years of age, however, most children noted that the speaker sometimes said happy things in a sad voice or vice versa. Children 6 years of age and younger maintained that the speaker did a good job of expressing her feelings; 9- and 10-year-olds expressed the opposite view. Although most 7- and 8-year-olds were aware that many of the test sentences contained conflicting emotional cues, they still maintained that the speaker had expressed her feelings well.

Discussion

The results indicate clear age-related changes in the relative allocation of attention to propositional and paralinguistic cues. Specifically, the majority of children 9 and younger judged the speaker's emotion by what she said; adults did so by how she spoke. The youngest children, 4-year-olds, showed the greatest reliance on propositional cues, almost all of them basing their judgments on what the speaker said. There were gradual, age-related increases in the use of paralinguistic cues to emotion, with roughly half of the 10-year-olds responding on the basis of paralanguage. In the early trials, participants of all ages took longer to respond to utterances with conflicting cues than to those with nonconflicting cues, indicating that they processed both sources of information. Nevertheless, children younger than 6 years of age failed to provide explicit descriptions of the conflicting cues, and those younger than 9 years of age expressed no reservations about the speaker's communication efficacy.

The results are consistent with claims of age-related continuities and discontinuities in affective reasoning. Continuity was evident in the processing of propositional and paralinguistic cues to emotion, as reflected in increased reaction times to messages with conflicting cues. Discontinuity was apparent in children's and adults' contrasting inferences about the speaker's feelings, and in younger and older children's descriptions of the emotionally ambiguous utterances.

An alternative explanation of the present findings is that the younger children were simply unable to label the paralinguistic cues that they detected. According to Dimitrovsky (1964), children 5 to 12 years of age become increasingly accurate at labeling affective paralanguage. Thus, their apparent preference for propositional cues in speech may stem from difficulties decoding affective paralanguage, an explanation that is at odds with greater response latency to messages with conflicting cues. Nevertheless, no firm conclusions about children's reasoning from conflicting cues can be drawn without clear indications that they can interpret the cues singly (Gnepp, 1983).

EXPERIMENT 2

The purpose of the second experiment was to evaluate the ability of the youngest participants in Experiment 1 (4-year-olds) to label affective paralanguage in stereotyped speech. To preclude any interference from message content, the test sentences were presented in a foreign language (Italian).

Method

Participants. The participants were 8 boys and 12 girls 4,0 to 4,3 years of age ($M = 4,1$), all from middle-class families who had volunteered to participate in campus research. An additional 2 boys and 3 girls were excluded from the data set because they were uninterested in the task.

Stimuli and apparatus. The stimuli consisted of two recorded versions of seven sentences produced by an Italian-speaking woman. In one version, she used stereotyped happy paralanguage; in the other, she used comparably sad paralanguage. The happy sentences had a higher average pitch level, faster speaking rate, and greater pitch and loudness variability than the sad sentences. All sentences were recorded digitally by means of SoundScope software and were presented by means of a Power Macintosh 8100 computer.

Procedure. Participants were told that they were going to play a listening game. They had to judge whether Rossana, the experimenter's friend, was feeling happy or sad. Although Rossana spoke a different language, they would know whether she was feeling happy or sad by listening carefully to her voice. Participants could respond by pointing to pictures of a happy or sad face or by saying "happy" or "sad" after each sentence. General noncontingent encouragement was provided.

Results

Of the 20 participants, 5 achieved a perfect score of 14, 6 achieved a score of 13, 7 achieved a score of 12, and 2 achieved a score of 11 correct. A binomial test, $n = 14$, maximum of 3 misses, $p = .8$, indicated that every child performed significantly above 80% correct, $p < .05$. Overall, children made 26 errors, 10 arising from the tendency for some children to begin by alternating between the two response options. When these alternation responses were removed from the analysis, performance was significantly better than 90% correct, binomial approximation to the norm, $n = 270$, hits = 254, $p\{\text{hit}\} = .9$, $p < .05$.

Discussion

Children 4 years of age accurately labeled the paralinguistic cues in Italian utterances, indicating that the decoding of paralanguage had not been the underlying problem in Experiment 1. For the 4-year-olds who exhibited difficulties, these difficulties were reflected in a response-alternation strategy. Jeffery and Cohen (1965) found that 3-year-olds tended to repeat the same response across trials in a two-choice task, in contrast to 4-year-olds, who alternated between the two responses. A number of children in the present experiment began by alternating between the two choices, but proceeded to respond appropriately after two or three trials. Verbal instructions proved unsuccessful in preventing this initial response tendency. In spite of this response bias in some children, performance was highly accurate.

Dimitrovsky (1964) also investigated the labeling of affective paralanguage by children 5 to 12 years of age. Although the 5-year-olds in her sample performed significantly above chance levels, they were considerably less accurate than the present 4-year-olds. Perhaps Dimitrovsky's use of four response alternatives instead of the two alternatives in the present task contributed to the differences. Moreover, the present study's use of an incomprehensible foreign language rather than English sentences with neutral content may have reduced potential distraction from verbal content. Finally, 12-year-olds' modest performance (66% correct) on Dimitrovsky's task implies that the paralinguistic cues in her test utterances were somewhat ambiguous.

The results are consistent with the notion that the children in Experiment 1 could interpret the affective paralanguage, but focused primarily on propositional content. Note, however, that the paralinguistic cues in the present experiment differed from those in Experiment 1. To provide unequivocal evidence that the paralinguistic cues in Experiment 1 were interpretable, this issue was examined directly.

EXPERIMENT 3

Although 4-year-olds could label the affective paralanguage in the Italian utterances of Experiment 2, it was necessary to demonstrate that they could also label the paralanguage in the utterances from Experiment 1. To remove potential distraction from the semantic content, the utterances were low-pass filtered. This procedure eliminates most phonetic cues, rendering the content unintelligible. Nevertheless, affective information transmitted by fundamental frequency and speaking rate is largely preserved (Rogers, Scherer, & Rosenthal, 1971; Scherer, Koivumaki, & Rosenthal, 1972).

Children in the present experiment were required to label the filtered materials as happy or sad by pointing to pictures of a happy or sad face or by providing verbal responses. Unlike the natural-sounding foreign sentences of Experiment 2 and the sentences with neutral content used by Dimitrovsky (1964), low-pass filtered speech sounds muffled. One consequence of such filtering is that variability from one sentence to the next is greatly reduced, making the task somewhat repetitive. Nevertheless, participants were expected to label the stimuli at better than chance levels.

Method

Participants. The participants were 60 children (thirty boys and thirty girls): twenty 4-year-olds (*range* =

4,0–4,3), twenty 5-year-olds (*range* = 5,0–5,3), and twenty 6-year-olds (*range* = 6,0–6,3). All participants were from middle-class families whose primary language was English. Data from an additional twenty-two children (twelve 4-year-olds, ten 5-year-olds) were excluded because of a response bias of alternating between happy and sad responses from trial to trial. Data from a further seven children were excluded because of failure to complete the task ($n = 2$), inattentiveness ($n = 4$), or reported problems with the headphones ($n = 1$).

Stimuli and apparatus. The stimuli were created by low-pass filtering the 40 digitally recorded sentences from Experiment 1 at 500 Hz (90 dB per octave rolloff). Pilot testing (13 adults) revealed that participants were at or near ceiling in labeling the sentences as happy or sad. Filtered sentences were then transferred to audiotape in two random orders. Of the 40 sentences, 8 (4 happy, 4 sad) were used as practice trials; the remaining 32 sentences were used in the test phase. The two tape orders were counterbalanced across age and gender. The materials were presented to children in a quiet room by means of a Sony TC-W32 Stereo Cassette Deck and Sony CD550 Digital Reference headphones. On the basis of pilot testing, the volume was adjusted to a comfortable listening level.

Procedure. Each child was tested individually. To familiarize children with the muffled quality of low-pass filtered speech, they listened first to three filtered sentences spoken with emotionally neutral paralanguage (i.e., neither happy nor sad). They were then given a brief orientation about happy and sad feelings, after which they received eight practice trials. Children were asked to listen to each sentence carefully, indicating whether the speaker sounded happy or sad. Children responded by pointing to one of two schematic faces (one happy, one sad) or by saying "happy" or "sad." After the practice trials, children received the 32 test trials. To help maintain interest, the experimenter provided general encouragement throughout the task. Testing took approximately 20 min for 6-year-olds and 30 min for 4- and 5-year-olds.

Results

Children 4, 5, and 6 years of age achieved mean correct scores (out of 32) of 20.6 ($SD = 5.8$), 26.3 ($SD = 5.1$), and 28.6 ($SD = 3.5$), respectively. A binomial approximation to the norm revealed that all age groups performed significantly above chance levels, $n = 640$, $p\{\text{hit}\} = .5$: for 4-year-olds, hits = 414, $z = 7.43$, $p < .01$; for 5-year-olds, hits = 525, $z = 16.2$, $p < .01$; for 6-year-olds, hits = 571, $z = 19.8$, $p < .01$. A two-way ANOVA (Age × Gender) revealed an effect for age, $F(2, 57) = 13.8$, $p < .01$, but no effect for gender and no Age × Gender interaction. Accordingly, data were collapsed across sex in a one-way ANOVA, yielding a main effect for age $F(2, 57) = 13.9$, $p < .01$. Post-hoc comparisons using Scheffé's procedure indicated that 4-year-olds performed more poorly than did 5-year-olds, mean difference = 5.7, $p < .01$, and 6-year-olds, mean difference = 8.0, $p < .01$, but the scores of 5- and 6-year-olds did not differ. The distribution of "happy" and "sad" responses did not differ across age groups, indicating that children did not show a "happy" or "sad" response bias.

Discussion

Children as young as 4 years of age were able to label the affective paralanguage of the low-pass filtered sentences, although accuracy improved with age. Whether older children's greater accuracy stemmed from better understanding of the paralinguistic information or from greater persistence on the repetitive task remains unclear. Nevertheless, the findings parallel those of Dimitrovsky (1964), except for considerably greater accuracy in the present experiment and the absence of improvement between 5 and 6 years of age.

The results from 4- and 5-year-olds potentially overestimate the ability of this age group because of the exclusion of twenty-two children (twelve 4-year-olds and ten 5-year-olds) who exhibited a response-alternation strategy. As noted, however, response alternation has been reported when 4-year-olds are unable to discriminate some stimuli in a two-choice task (Jeffery & Cohen, 1965). Although those excluded from the final sample of the present experiment may have been unable to discriminate the stimuli, it is more likely that they were unmotivated to participate in a task that most children found boring. Informal observations, as recorded in notes from each test session, indicated that very shy children tended to adopt a response-alternation strategy along with a disinclination to communicate with the experimenter. For the most part, however, 4-, 5-, and 6-year-old children were able to label the content-filtered sentences from Experiment 1 as either happy or sad. In conclusion, it is unlikely that the response patterns in Experiment 1 were attributable to children's inability to label the paralinguistic information. Instead, they probably arose from children's greater attention to propositional content.

GENERAL DISCUSSION

In Experiment 1, 4- to 10-year-olds and adults listened to utterances with conflicting or nonconflicting

propositional and paralinguistic cues to happiness and sadness. When the cues to emotion conflicted, children 8 years of age and younger judged the speaker's feelings by what she said (i.e., the propositional content), whereas adults judged the speaker's feelings by how she spoke (i.e., the paralanguage); 9- and 10-year-olds were divided in their propositional or paralinguistic focus. In the early trials, listeners of all age groups hesitated when responding to messages with conflicting cues, which confirmed their processing of paralinguistic and propositional cues. When queried about unusual aspects of the messages, only participants 7 years of age and older were consistent in reporting that the speaker said happy things in a sad way (or vice versa). By contrast, 4- to 6-year-olds commended the speaker's efficacy in communicating her feelings.

Young children's focus on content cannot be attributed to interpretive difficulties with paralanguage, as indicated by their accurate decoding of comparable paralanguage in Experiments 2 and 3 when the content was obscured (i.e., by use of a foreign language and low-pass filtered utterances) and by their greater response latency to messages with conflicting cues. In other words, children could interpret the paralinguistic cues in the original sentences, but they simply accorded greater weight to the propositional cues.

The developmental shift in the relative salience of propositional and paralinguistic cues to emotion parallels the reported shift in interpreting comparable cues to speaker attitude (Bugental et al., 1970; Friend & Bryant, 2000; Mehrabian & Wiener, 1967; Solomon & Ali, 1972). The unique contribution of the present study, however, is its unequivocal demonstration that children and adults base their judgment of a speaker's happiness or sadness on very different grounds. For children in the present study, the critical cues concerned the positive or negative emotional implications of events depicted by the speaker. This propositional focus prevailed despite the absence of specific emotion words (e.g., *happy, sad*) or emotionally charged words (e.g., *fantastic*, *disappointing*), which might have accounted for the propositional bias observed in previous studies (e.g., Friend & Bryant, 2000). For adults, vocal affect seemed to function as a barometer of the speaker's feelings. Why, then, were children and adults so different in this respect?

One possibility is that young children failed to understand the experimenter's instructions, which directed them to listen to the speaker's voice and indicate whether she was feeling happy or sad. If young children confused *message* (i.e., what was said) with *voice* (vocal tone or quality), their focus on content would be understandable. Alternatively, they could have ignored the word *voice* because of its incomprehensibility, in which case they would have no basis for altering their habitual focus on content. This type of explanation, while potentially applicable to the younger children in the present sample, is clearly inapplicable to the many 8- to 10-year-olds who also focused on content. In fact, children 6 years of age and older often used the term *voice* in their explanations of the speaker's "silly" way of talking (i.e., saying happy things in a sad voice).

Another possibility is that children have limited understanding of the role of vocal emotion in communication (Levy, 1982). As a result, they fail to consider its relevance to the speaker's feelings. Despite their ability to decode paralinguistic cues, children may not treat such cues as a basis for qualifying or even overriding the propositional content or literal message. Explanations such as these can account for children's difficulty with irony (Ackerman, 1986) and sarcasm (Capelli, Nakagawa, & Madden, 1990), both of which require the integration of contextual or paralinguistic cues with opposing literal meanings. The task in the present study was considerably easier in that it did not require any integration of propositional and paralinguistic information. Instead, the response options, happy and sad, demanded a focus on one type of cue, propositional or paralinguistic.

Although it might be tempting to invoke notions such as attentional flexibility (e.g., Bialystok, 1993; Zelazo, Frye, & Rapus, 1996) or the inhibition of habitual responses (Carlson, Moses, & Hix, 1998; Diamond & Taylor, 1996) to account for the adult–child differences, such explanations are clearly inappropriate for the older children. Indeed, more 9-year-olds focused on content than on paralanguage, as did many 10-year-olds. Moreover, the majority of children 6 years of age and older provided clear descriptions of the conflicting propositional and paralinguistic cues, which confirmed their awareness of both sources of information. Regardless of the underlying reasons, it is clear that children as old as 10 years of age accord considerably less weight to vocal cues to emotion than do adults. There are indications, however, that facial cues can override the happy or sad implications of events depicted visually (Gnepp, 1983). Whether such facial cues can override contradictory verbal cues is an important question for future research.

Despite infants'sensitivity to paralanguage (Dimitrovsky, 1964; Fernald, 1993; Papoušek et al., 1990; Walker-Andrews & Grolnick, 1983) and toddlers' comprehension and production of emotion words (Borke, 1971; Bretherton & Beeghly, 1982; Bretherton et al., 1986; Izard, 1971; Ridgeway et al., 1985), much

remains unclear about preschool- and school-age children's understanding of vocal expressiveness. Nevertheless, the present findings indicate that children and adults reach opposite conclusions about a speaker's feelings in the context of mismatched propositional and paralinguistic cues to happiness and sadness. Whether the presence of contradictory cues within a single modality contributes to these differences remains to be determined. In any case, it is of particular importance to ascertain the factors responsible for the pattern of findings observed.

ACKNOWLEDGMENTS

Funding for this research was provided by a grant from the Natural Sciences and Engineering Research Council of Canada to Sandra E. Trehub and an Ontario graduate scholarship to J. Bruce Morton. The authors are grateful to Marianne Fallon and Rossana Bisceglia for assistance with stimulus preparation, and to Gary Fernandes and Salima Ibrahim for assistance in the testing of children. Thanks also go to Phil Zelazo and Meredyth Daneman for helpful comments on an earlier version of the manuscript.

ADDRESSES AND AFFILIATIONS

Corresponding author: Sandra E. Trehub, Department of Psychology, University of Toronto at Mississauga, Mississauga, ON L5L 1C6 Canada; e-mail: sandra.trehub@utoronto.ca. J. Bruce Morton is at the University of Denver, Colorado.

REFERENCES

Ackerman, B. P. (1986). Children's sensitivity to comprehension failure in interpreting a nonliteral use of an utterance. *Child Development, 57*, 485–497.

Bachorowski, J. A., & Owren, M. J. (1995). Vocal expression of emotion: Acoustical properties of speech are associated with emotional intensity and context. *Psychological Science, 6*, 219–224.

Backscheider, A. G., & Gelman, S. A. (1995). Children's understanding of homonyms. *Journal of Child Language, 22*, 107–127.

Bamberg, M., & Reilly, J. (1996). Emotion, narrative, and affect: How children discover the relationship between what to say and how to say it. In D. I. Slobin, J. Gerhardt, A. Kyratzis, & J. Guo (Eds.), *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp* (pp. 329–341). Mahwah, NJ: Erlbaum.

Beal, C. R., & Flavell, J. H. (1982). Effect of increasing the salience of message ambiguities on kindergartners' evaluations of communicative success and message adequacy. *Developmental Psychology, 18*, 43–48.

Bearison, D. J., & Levey, L. M. (1977). Children's comprehension of referential communication: Decoding ambiguous messages. *Child Development, 48*, 716–720.

Bialystok, E. (1993). Metalinguistic awareness: The development of children's representations of language. In C. Pratt & A. F. Garton (Eds.), *Systems of representation in children: Development and use* (pp. 211–233). Chichester, U.K.: Wiley.

Bloom, L. (1993). *The transition from infancy to language: Acquiring the power of expression.* New York: Cambridge University Press.

Bloom, L., & Beckwith, R. (1989). Talking with feeling: Integrating affective and linguistic expression in early language development. *Cognition and Emotion, 3*, 313–342.

Bolinger, D. L. (1986). *Intonation and its parts.* Stanford, CA: Stanford University Press.

Borke, H. (1971). Interpersonal perception of young children: Egocentrism or empathy? *Developmental Psychology, 5*, 263–269.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology, 18*, 906–921.

Bretherton, I., Fritz, J., Zahn-Waxler, C., & Ridgeway, D. (1986). Learning to talk about emotions: A functionalist perspective. *Child Development, 57*, 529–548.

Bugental, D. E., Kaswan, J. W., & Love, L. R. (1970). Perception of contradictory meanings conveyed by verbal and nonverbal channels. *Journal of Personality and Social Psychology, 16*, 647–655.

Burns, N., & Cavey, L. (1957). Age differences in empathic ability among children. *Canadian Journal of Psychology, 11*, 227–230.

Capelli, C. A., Nakagawa, N., & Madden, C. M. (1990). How children understand sarcasm: The role of context and intonation. *Child Development, 61*, 1824–1841.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development, 69*, 672–691.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584–1595.

Cramer, P. (1983). Homonym understanding and conservation. *Journal of Experimental Child Psychology, 36*, 179–195.

Deutsch, F. (1974). Female preschoolers' perceptions of affective responses and interpersonal behavior in videotaped episodes. *Developmental Psychology, 10*, 733–740.

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do." *Developmental Psychobiology, 29*, 315–334.

Dimitrovsky, L. (1964). The ability to identify the emotional meaning of vocal expression at successive age levels. In J. Davitz (Ed.), *The communication of emotional meaning* (pp. 69–86). New York: McGraw-Hill.

Dunn, J., Brown, J., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology, 27*, 448–455.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195.

Fernald, A. (1991). Prosody in speech to children: Prelinguistic and linguistic functions. *Annals of Child Development, 8*, 43–80.

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). New York: Oxford University Press.

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64*, 657–674.

Frick, R.W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97*, 412–429.

Friend, M., & Bryant, J. B. (2000). A developmental lexical bias in the interpretation of discrepant messages. *Merrill-Palmer Quarterly, 46*, 342–369.

Gnepp, J. (1983). Children's social sensitivity: Inferring emotions from conflicting cues. *Developmental Psychology, 19*, 805–814.

Greenspan, S., Barenboim, C., & Chandler, M. J. (1976). Empathy and pseudo-empathy: The affective judgments of first- and third-graders. *Journal of Genetic Psychology, 129*, 77–88.

Hoffner, C., & Badzinski, D. M. (1989). Children's integration of facial and situational cues to emotion. *Child Development, 60*, 411–422.

Izard, C. E. (1971). *The face of emotion.* New York: Appleton-Century-Crofts.

Jeffery, W. E., & Cohen, L. B. (1965). Effect of spatial separation of stimulus, response, and reinforcement in selective learning in children. *Journal of Experimental Psychology, 67*, 577–580.

Jerger, S., Martin, R. C., & Pirozzolo, F. J. (1988). A developmental study of the auditory Stroop effect. *Brain and Language, 35*, 86–104.

Jusczyk, P. W. (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.

Levy, R. I. (1982). On the nature and functions of the emotions: An anthropological perspective. *Social Science Information, 21*, 511–528.

Levy, R. I. (1984). The emotions in comparative perspective. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 397–412). Hillsdale, NJ: Erlbaum.

Lewis, M. M. (1951). *Infant speech.* London: Routledge & Kegan Paul.

Mehrabian, A., & Wiener, M. (1967). Decoding of inconsistent communications. *Journal of Personality and Social Psychology, 6*, 109–114.

Papoušek, M., Bornstein, H., Nuzzo, C., Papoušek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development, 13*, 539–545.

Reilly, S. S., & Muzekari, L. H. (1979). Responses of normal and disturbed adults and children to mixed messages. *Journal of Abnormal Psychology, 88*, 203–208.

Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology, 21*, 901–908.

Rogers, P. L., Scherer, K. R., & Rosenthal, R. (1971). Content-filtering human speech. *Behavioral Research Methods and Instrumentation, 3*, 16–18.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99*, 143–165.

Scherer, K. R., Koivumaki, J., & Rosenthal, R. (1972). Minimal cues in the vocal communication of affect: Judging emotions from content-masked speech. *Journal of Psycholinguistic Research, 1*, 57–64.

Solomon, D., & Ali, F. A. (1972). Age trends in the perception of verbal reinforcers. *Developmental Psychology, 7*, 238–243.

Solomon, D., & Yaeger, J. (1969). Determinants of boys' perceptions of verbal reinforcers. *Developmental Psychology, 1*, 637–645.

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*, 188–195.

Walker-Andrews, A., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior and Development, 6*, 491–498.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43*, 230–246.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37–63.