HHS Public Access
Author manuscript
*Merrill Palmer Q (Wayne State Univ Press)*. Author manuscript; available in PMC 2017 July 09.

Published in final edited form as:
*Merrill Palmer Q (Wayne State Univ Press)*. 2000 April ; 46(2): 342–369.

# A Developmental Lexical Bias in the Interpretation of Discrepant Messages

**Margaret Friend** and
Department of Psychology, University of South Florida and Department of Cognitive Sciences, University of California, Irvine

**Judith Becker Bryant**
Department of Psychology, University of South Florida

## Abstract

Children's interpretations of lexical and vocal cues to speaker affect, independently and in combination, were examined in four studies. In Experiments 1 and 2, 7- to 11 -year-olds'judgments of lexical and paralinguistic cues were evaluated. In Experiment 3, these cues were combined to produce consistent and discrepant messages. The affective interpretations of 7- to 10-year-olds reflected a weighted-averaging strategy favoring the affect conveyed lexically. In Experiment 4, the developmental trajectory of children's interpretations of discrepancy from 4 to 10 years of age was investigated. Both 4- and 7-year-olds appeared to use a weighted-averaging strategy favoring lexical content, whereas 10-year-olds utilized a strategy favoring paralanguage.

The focus of the present paper is a developmental phenomenon that has implications for cognitive-developmental theory: children's tendency to be biased toward lexical interpretations of affectively discrepant auditory messages (e.g., sarcasm and joking). Social-referencing studies reveal an important role for vocal expression (paralanguage) in infant behavior regulation (Mumme, Fernald, & Herrera, 1996; Svejda, 1982). Yet rating scale studies suggest that children of preschool and early school age give less weight to vocal paralanguage than to lexical content in interpreting affective discrepancy (Friend, in press; Reilly & Muzekari, 1986; Solomon & Ali, 1972; cf. Bugental, Kaswan, & Love, 1970). For example, a speaker saying "You're my favorite person" in an angry voice would be interpreted as feeling happy. Similar effects are obtained for negative lexical content spoken in a happy, approving, or positive voice (Bugental et al., 1970; Friend, in press; Solomon & Ali, 1972). And some authors have reported that this effect is exaggerated among children referred for psychological disturbance (Reilly & Muzekari, 1986). In contrast, adults appear to base affective judgments of sarcastic or joking utterances on paralinguistic, rather than lexical, content (Argyle, Alkema, & Gilmour, 1971; Mehrabian & Wiener, 1967; Reilly & Muzekari, 1986).

Children's understanding of sarcasm also has been investigated using a story-telling paradigm to contrast contextual and paralinguistic cues. Winner et al. (1987) found that

Correspondence may be sent to Margaret Friend, who is now at the Department of Psychology, San Diego State University, 6363 Alvarado Court, Suite 103, San Diego, California 92120. Electronic mail may be sent via Internet to mfriend@sunstroke.sdsu.edu.

contextual cues facilitated 8-year-olds' recognition of sarcasm, whereas paralinguistic cues did not. In contrast, Capelli, Nakagawa, and Madden (1990) found that paralinguistic cues facilitated the recognition of sarcasm to a greater extent than contextual cues. In neither study were the independent contributions of lexical and paralinguistic content to the perception of discrepancy assessed, making these divergent findings difficult to resolve. However, the weight of the evidence across paradigms suggests that children have difficulty with paralinguistic cues when they are paired with discrepant lexical content. More recently, Dews et al. (1996) provided evidence that, as early as 6 years of age, children have at least a rudimentary understanding of discrepant communication, but they did not assess the relative weighting of message components.

Three tasks are critical to the elucidation of lexical bias in development and motivate the present research. First, it is necessary to assess sensitivity to lexical and paralinguistic cues both in isolation and in combination in the 7- to 10-year-old age group in which lexical bias has been reported (Bugental et al., 1970; Friend, in press; McCluskey & Albas, 1981; Reilly & Muzekari, 1979,1986; Solomon & Ali, 1972). Importantly, this is also the age group addressed in recent cognitive-developmental research (Capelli et al., 1990; Dews et al., 1996; Winner et al., 1987). Second, assessment of the generalizability of lexical bias to children who have experience with discrepancy (Blotcky, Tittler, & Friedman, 1982; Bugental, Love, Kaswan, & April, 1971) is a first step toward understanding the ubiquitousness of lexical bias and its amenability to training. Finally, programmatic research, grounded in cognitive-developmental theory, is required to reconcile the appearance of a lexical bias in child-hood with theoretical approaches emphasizing the primacy of the voice in communicating affective information (Scherer, 1991).

Conceptually, the presence of a developmental lexical bias in affective judgments presents a paradox. Vocal paralanguage has long been considered a profound source of affective information (Darwin, 1965/ 1873). According to Scherer (1991, p. 146), "Affect vocalizations are the closest we can get to the pure biological expression of emotion and one of the most rudimentary forms of communication" The significance of vocal expression for infant behavior regulation lends credence to this position (Mumme et al., 1996; Svejda, 1982). Yet, in childhood, vocal paralanguage appears relegated to a secondary status relative to words that appear to be more salient in making inferences regarding a speaker's affect. Research to date has focused on the transition in middle childhood toward an increased reliance on paralinguistic cues in affective judgments (Friend, in press; Solomon & Ali, 1972).

Early accounts of children's utilization of cues to speaker affect share, implicitly, the assumption that experience with resolving discrepancy promotes more adultlike performance (Bugental et al., 1970, Solomon & Ali, 1972). It follows, then, that children who have extensive experience with affective discrepancy should exhibit a tendency to interpret speaker affect in accord with paralinguistic cues (Argyle et al., 1971; Mehrabian & Wiener, 1967; Reilly & Muzekari, 1986). However, comparisons of disturbed children (who have more extensive experience with discrepancy than their nondisturbed siblings and peers; Blotcky et al., 1982; Bugental et al., 1971) and nondisturbed children have not yielded support for this assumption. Some researchers report no difference in the interpretation of discrepancy across disturbed and nondisturbed children, and others report an exaggerated tendency to base interpretations on lexical cues among disturbed children (McCluskey & Albas, 1981; Reilly & Muzekari, 1979, 1986). This is not to suggest (vs. Bateson, Jackson, Haley, & Weakland, 1956) that discrepant communication is causal in specific pathological outcomes. Rather, the comparison of disturbed and nondisturbed children permits an indirect assessment of whether experience with discrepancy influences the relative weighting of lexical and paralinguistic cues. It also provides a sense of the extent to which a lexical bias in affective interpretations is characteristic of both normal and disordered development.

McCluskey and Albas (1981) examined boys' perceptions of affectively discrepant auditory messages. Boys with emotional disturbance reported feeling more negatively in response to discrepancy, but their interpretations did not differ from those of their nondisturbed peers. In two experiments, Reilly and Muzekari (1979, 1986) examined the weighting of lexical and paralinguistic cues in samples with and without emotional disturbance. Consistent with previous findings (Argyle et al., 1971; Mehrabian & Wiener, 1967) nondisturbed adults interpreted discrepant messages on the basis of paralinguistic cues. However, children between 7 and 9 years of age interpreted these messages in terms of lexical content. In contrast to McCluskey and Albas's (1981) findings, the tendency to base interpretations on lexical cues was most pronounced among children referred for emotional disturbance.

Two issues limit the generalizability of these findings. First, judgment studies of paralanguage (in the absence of lexical content) were not conducted, leaving the relative salience of lexical and paralinguistic cues indeterminate. Second, it is not clear whether the source of individual differences in the Reilly and Muzekari studies (1979, 1986) is disturbance per se or other demographic factors such as socioeconomic status (SES), ethnicity, or IQ, which may be associated with psychological referral.

In the present research lexical bias in affective judgments was examined using both a methodologically refined individual differences approach in which we address directly the limitations of previous research in this area (Experiments 1–3) and a cross-sectional approach in which we evaluate the developmental trajectory of children's discrepancy resolution strategies (Experiment 4). In Experiments 1 and 2, independent ratings were obtained of both lexical and paralinguistic content using adult raters. We then assessed the accuracy and interrater agreement of children's judgments of lexical and paralinguistic items rated at criterion by adults. Experiments 1 and 2 served two purposes: the selection, for presentation in Experiment 3, of a set of items whose lexical and paralinguistic content were reliably rated by adults, and the assessment of child accuracy and agreement on lexical and paralinguistic cues presented in isolation.

In Experiment 3, the composition of samples with and without emotional disturbance was roughly comparable with respect to age, SES, ethnicity, gender, and estimated IQ. The generalization of lexical bias across these groups has implications for the role of experiential mechanisms in children's developing understanding of affective discrepancy. Illustratively, there are three possible outcomes: First, the presence of a lexical bias only among participants with disturbance would necessitate an account based solely on differential developmental influences (e.g., differential rates of experience with discrepancy). Second, the generalization of lexical bias across groups would necessitate an account based on general influences common to both groups (e.g., developmental changes in processing efficiency). Finally, evidence of a lexical bias across groups that varies as a function of disturbance would necessitate an account based on both general and differential mechanisms.

In Experiment 4, developmental changes in children's weighting of lexical and paralinguistic information from age 4 to age 10 were assessed. An important focus was the generalization of effects obtained in Experiment 3 to a new sample of children using a new set of experimental stimuli. A second focus was the empirical estimation of the developmental transition in children's weighting strategies from those favoring lexical content to more adultlike strategies favoring paralanguage.

# EXPERIMENT 1

## Method

**Participants—**Adult under-graduate psychology students who volunteered to participate. Child raters were 12 children between 7 and 10 years of age (*M* = 9;3) selected from the county school system. Children were roughly matched with children in Experiment 3 on ethnicity, gender, and SES (operationally defined as eligibility for school lunch programs, an indicator of SES in studies of academic performance [Carswell & White, 1984; Marcon, 1993]). The demographic characteristics of this sample were as follows: low SES Black males, 25%; low SES Black females, 8%; higher SES Black females, 8%; low SES White males, 25%; low SES White females, 8%; higher SES White males, 16%; higher SES White females, 8%. Children were screened for reading ability on the basis of information provided by their teachers. The mean estimated IQ for this group was 87.5 (*SD* = 11.2).

**Stimuli—**Fifty-four sentences were selected from corpora of mother-child interactions obtained by D. A. Bugental (personal communication, October 17, 1984) and the second author. Of these, 27 sentences were selected to convey happiness (e.g., "You're doing a great job") and anger (e.g., "You're nothing more than a big baby").

Typescripts of 54 sentences were rated for their affective content by adult raters using a 5-point Likert scale with points labeled *really happy, sort of happy, just okay, sort of mad*, and *really mad*. Sentences were listed in random or reversed order (counterbalanced across raters) with the restriction that no more than three sentences representing the same affect were listed consecutively. Interrater agreement was calculated for each sentence using the formula: agreement = # of agreements/ # of agreements plus # of disagreements. Criterion agreement was .80 in the direction of the intended affective state (happy or angry). Due to an effect of central tendency, the scale was collapsed so that a rating of 1 or 2 was scored as happy, and a rating of 4 or 5 was scored as angry. Adult agreement met or exceeded criterion for 30 of the 54 stimulus sentences. Of these, 12 were happy and 18 were angry. Six angry sentences were randomly dropped from the study in order to present an equal number of happy and angry (12 each) sentences to children.

**Procedure—**Typescripts of 24 stimulus sentences, rated at .80 agreement by adults were presented individually to children in a room at their school. Instructions were read aloud but stimulus sentences were presented in written form only. Order of presentation (random or reversed) was counterbalanced across raters. Children used a 5-point Likert scale with faces drawn above each point labeled *really happy, sort of happy, just okay, sort of mad*, and *really mad*. The scale is similar to one designed by Buck (1975), and pilot testing confirmed its utility. A short form WISC-R administered after completion of the rating task yielded estimated IQs. The short form WISC-R consists of vocabulary and block design subscales and correlates at .91 with full scale IQ (Sattler, 1982).

### Results

Chi-square one-variable goodness-of-fit tests indicated that children's judgments were significantly more accurate than chance for all 24 stimulus sentences, $\chi^2(2, N = 12)$, $p < .05$ (see Table 1 for means). Children's interrater agreement reached .80 for 19 sentences (12 happy and 7 angry) or 79% of those rated at criterion by adults. All 19 sentences were included in Experiment 2.

## EXPERIMENT 2

### Method

**Participants—**Adult raters were three male and three female under-graduate psychology students who volunteered to participate. Child raters were 12 children, 7 to 11 years of age ($M$ = 9;5), selected from the county school system. Children were screened for hearing impairment on the basis of information provided by their principal and were completely matched with children in Experiment 1 on ethnicity, gender, and SES. The mean estimated IQ for this sample was 88.58 ($SD$ = 13).

**Stimuli—**Most previous attempts to isolate vocal paralanguage from lexical content have relied on low-pass filtering (Rogers, Scherer, & Rosenthal, 1971). Low-pass filtering preserves the fundamental frequency contour ($F_0$; perceived as pitch) but degrades intensity (perceived as loudness) and information contained in the harmonic structure (which contributes to the percept of voice quality). An alternative technique, reiterant speech, involves replacing the original syllables of an utterance with nonsense syllables to produce an utterance as acoustically similar as the original utterance as possible (Liberman & Streeter, 1978; Nakatani & Schaffer, 1978). Raters' judgment accuracy tends to be better for reiterant stimuli than for low-pass filtered stimuli. In practical terms, this means that a larger number of stimuli are rated at criterion during stimulus validation and can be retained for subsequent experiments (Friend & Farrar, 1994). Also, children find reiterant stimuli easier to rate than low-pass filtered stimuli (Friend, in press). For these reasons, reiterant speech was used to isolate the paralinguistic content of messages in the present research.

The experimenter, a female native speaker of American English, produced the stimulus messages. Messages were recorded onto Maxell reel-to-reel tape using a TEAC Model 3340S laboratory tape recorder. Acoustic description of speech samples was obtained using a Visi-Pitch Model 6087 pitch-extracting device.

To produce reiterant stimuli, a single syllable ("ma") was repeated to mimic the paralinguistic information that was used in the original utterance. The reiterant speech version of the sentence "You're doing a great job!,"for example, was read as: "Ma MaMa Ma Ma Ma!" Determination of the acoustic similarity of original and reiterant stimuli was based on their respective $F_0$ contours and $F_0$ ranges (the difference between the highest and lowest fundamental frequencies). Only those reiterant stimuli from original-reiterant stimulus pairs with visually similar $F_0$ contours and $F_0$ ranges that differed by less than 15 Hz were used in this experiment.

The experimenter read each sentence in both happy and angry voice (original stimuli) and produced reiterant versions for each. Thirty-eight original (for use in Experiment 3) and reiterant stimuli (for validation purposes in Experiment 2) were recorded: 12 happy sentences (in happy and angry paralanguage) for a total of 24 original-reiterant pairs, and 7 angry sentences (in happy and angry paralanguage) for a total of 14 original-reiterant pairs. Thirty-eight reiterant stimuli (19 in happy paralanguage and 19 in angry paralanguage) were recorded onto cassette tape using a Technics Model RS-B14 cassette deck. Stimulus order was randomized with the restriction that no more than three samples of the same affect occurred consecutively.

Acoustic measures are reported as an indication of the similarity of original (paralinguistic and lexical content) and reiterant (paralinguistic content only) stimuli. The difference in mean $F_0$ between original and reiterant stimuli was small ($M$ = 7.1 Hz, $SD$= 4 Hz) across all stimuli. The difference was *7.7* Hz ($SD$ = 4 Hz) for happy stimuli and 6.4 Hz ($SD$ = 3.9 Hz) for angry stimuli.

The mean $F^0$ range was nearly identical for original ($M$ range = 136.7 Hz, $SD$ = 41.4 Hz) and reiterant stimuli ($M$ range = 134.9 Hz, $SD$ = 42.6 Hz). Importantly, $F_0$ range contributes more to the perception of specific affective states than does mean $F_0$ (Ladd, Silverman, Tolkmitt, Bergman, & Scherer, 1985). Substantial differences in $F_0$ range were obtained between happy paralanguage and angry paralanguage stimuli (see Figures 1 and 2). Happy stimuli yielded a mean $F_0$ range of 261.5 Hz ($SD$ = 49.7 Hz), whereas the same lexical content read in an angry voice yielded a mean $F_0$ range of 133.3 Hz ($SD$= 44 Hz). Perceptually, this translates into greater variation in pitch for utterances with happy, relative to angry, paralanguage.

The 38 reiterant stimuli were played individually to adult raters on an AIWA Model HS-P02 cassette player. Stimulus affect was rated on a 5-point scale identical to the one used in Experiment 1. Interrater agreement was computed on a collapsed scale as in Experiment 1. Criterion agreement was .80 in the direction of the intended affect.

Criterion agreement was met for 22 stimuli (13 happy and 9 angry). Because we wanted to present the same lexical content across consistent and discrepant conditions, we selected those stimuli for which both happy and angry reiterant versions were rated at criterion (4 happy and 3 angry lexical content). Two additional sentences (1 happy and 1 angry) for which criterion was reached for the happy reiterant versions only were included. Agreement for the angry versions of these stimuli was .67 (four out of six raters). The resulting stimulus set included 9 sentences (5 happy and 4 angry lexical content) produced in happy and angry paralanguage for a total of 18 reiterant stimuli (9 happy and 9 angry paralanguage).

**Procedure—**The 18 reiterant stimuli validated by adults were presented in random and reversed order (counterbalanced across participants) over a set of headphones. Children completed the rating task individually in a room at school using a 5-point scale identical to the one in Experiment 1. After the rating task, a short form WISC-R was administered.

## Results

Children correctly identified the affect of 67% (or 12) of the 18 stimuli at a level exceeding chance, $\chi^2(2, N = 12)$, $p < .05$, a figure consistent with the literature on children's detection of vocal paralanguage in the absence of lexical content (Baltaxe, 1991; Dimitrovsky, 1964; Matsumoto & Kishimoto, 1983). The mean rating was 2.08 ($SD = 0.66$) for happy and 3.77 ($SD = 0.68$) for angry reiterant stimuli. This finding suggests that children in this age range utilize paralinguistic cues to speaker affect.

The significant chi-squares reflect an interrater agreement of at least .67 among child raters on the majority of the reiterant stimuli. Thus, utilizing a criterion interrater agreement of .67 for child raters results in retention of the majority of stimuli (12 of 18). However, children reached adult interrater agreement levels (.80) for only 22% (or 4) of 18 stimuli. Consideration of both the accuracy and agreement data from Experiments 1 and 2 suggests that children are, on the average, accurate judges of lexical content and paralanguage but that there is greater variability in para-linguistic than in lexical judgments.

Affective intensity (a measure of the relative salience of lexical and paralinguistic cues) was compared across Experiments 1 and 2 for adults and children. It was operationally defined as the point on the rating scale assigned to a stimulus. For example, a happy reiterant stimulus (para-language) and a happy sentence (lexical content) each receiving a rating of 1 (really happy) were considered relatively equivalent (across the separate groups of raters in Experiments 1 and 2) in affective intensity. Mean ratings of lexical and paralinguistic content as a function of rater (adult or child) and intended affect (happy or angry) were compared. Bonferroni dependent samples $t$-tests ($fw\ \alpha = .05$) were conducted to determine whether differences in lexical and paralinguistic intensity reached significance (see Table 1 for means and standard deviations).

There were no significant differences in lexical and paralinguistic intensity for adult raters. However, children (in Experiment 2) rated para-linguistic content as less intense, on the average, than children (in Experiment 1) rated lexical content. This is not surprising given that children's interrater agreement was lower on paralinguistic than on lexical cues. Children may perceive paralinguistic cues as more ambiguous than lexical cues even though adults rate the same items as quite unambiguous. This is an important factor to consider, and one we will return to, when evaluating children's interpretations of full sentences (lexical and paralinguistic content) in Experiment 3.

## EXPERIMENT 3

Experiments 1 and 2 suggest that children are more consistent judges of lexical than of paralinguistic cues, even though they are more accurate than chance in both cases. Still at issue is whether children's judgments of lexical content and paralanguage in Isolation presage a lexical bias in their affective ratings of lexical/paralinguistic discrepancies. Further, the assumption that experience with complex communications facilitates interpretations based on paralanguage remains to be assessed. The comparison of children referred for psychological disturbance with nonreferred children provides a preliminary evaluation of this assumption.

### Method

**Participants—**Thirty-four children, 7 to 10 years of age (*M* = 8;9), were selected from regular classrooms in the county school system. The children were screened for hearing impairment and were roughly matched on ethnicity, gender, and SES with the disturbed sample. The demographic characteristics of this sample were as follows: low SES Black males, 21 %; low SES Black females, 3%; higher SES Black females, 15%; low SES White males, 15%; low SES White females, 21%; higher SES White males, 9%; and higher SES White females, 16%. The mean estimated IQ was 93.5 (*SD* = 13.5).

Thirty-four children, 7 to 10 years of age (*M* = 8;9), were chosen from the county's SED (Severely Emotionally Disturbed) Program. Children in this program exhibit chronic behavior problems that interfere with the learning process and are not attributable to physical, sensory, or intellectual deficits. It is the most extreme classification within the school system and refers to a range of psychiatric disorders. Children were screened for hearing impairment and roughly matched on ethnicity, gender, SES, and age with the children in the previous sample. The demographic characteristics of this sample were as follows: low SES Black males, 26%; low SES Black females, 3%; higher SES Black females, 9%; low SES Hispanic males, 3%; low SES White males, 15%; low SES White females, 24%; higher SES White males, 14%; and higher SES White females, 6%. The mean estimated IQ was 88.1 (*SD* = 12).

**Stimuli—**All original, full sentence (lexical and paralinguistic content), stimuli corresponding to the reiterant stimuli rated at criterion by adults in Experiment 2 were retained. Stimuli were recorded onto cassette tape in random and reversed orders with the restriction that no more than two sentences representing the same affect were recorded consecutively. Each experimental order included 18 stimuli across four stimulus types: 5 happy lexical content-happy paralanguage, 5 happy lexical content-angry para-language, 4 angry lexical content-angry paralanguage, and 4 angry lexical content-happy paralanguage.

**Procedure—**A toy store scenario in which toys and shelves block a child's view, but the child can overhear a "mom" talking, was described in order to suggest a context in which a child might receive only auditory information about a speaker. Children were told that they would be answering a question about what they heard. This procedure was implemented by two experimenters each testing equal proportions of participants from each sample. The experimenter showed the child the rating scale used in Experiments 1 and 2, labeled each face verbally, then asked the child to label each face. All children accurately labeled the faces on the rating scale. Children heard 18 full sentence stimuli presented individually in a room at the child's school in either random or reversed order (counterbalanced within groups) over headphones. After each stimulus, the child was asked, "How do you think this mom is feeling?" Children responded by pointing to a face on the rating scale. Following the rating task, a short form WISC-R was administered.

## Results

Planned comparisons indicated that the difference in mean estimated IQ between samples was nonsignificant, so standard multivariate analyses were appropriate. Given that the groups also were comparable on gender, ethnicity, SES, and age, this suggests that the primary difference between the experimental samples was referral to the school system's SED program.

Three sets of analyses were conducted. First, a Disturbance (2) × Age (2) × Paralanguage (2) × Discrepancy (2) mixed models repeated-measures MANOVA was conducted on children's affective ratings of the full stimulus set. Next, item level performance was assessed to determine the extent to which effects generalized across individual stimuli. Finally, discrepancy resolution strategies were evaluated for individual participants in both experimental samples to determine the extent to which the average data reflected individual performance. The omnibus analysis was conducted at a = .05. For all subsequent comparisons, $fw\alpha = .05$.

Before conducting the omnibus analysis, a median split was performed on children's age in months across disturbance classification for the purpose of exploring developmental effects (median age = 102.5 months). This produced two arbitrary age groups within each sample. For children without emotional disturbance, the mean age of the younger group was 93 months (range = 86–102 months, N= 20) and the mean age of the older group was 116 months (range = 103–131 months, $N = 14$). For children with emotional disturbance, the mean age of the younger group was 90 months (range = 79–96 months, $N = 12$) and the mean age of the older group was 113 months (range = 103–130 months, $N = 22$). In this way it was possible to consider whether younger and older children differ in their interpretations of discrepancy.

The MANOVA yielded main effects of paralanguage and disturbance and two significant interactions: Paralanguage × Discrepancy and Paralanguage × Discrepancy × Disturbance (see Tables 2 and 3 for means and standard deviations). The analysis yielded no effects of age, and the same pattern of findings was obtained when the age variable was excluded from the analysis. The main effect of paralanguage indicated that, across levels of discrepancy, angry paralanguage messages were rated as more angry ($M = 3.26$) than happy paralanguage messages, $M = 2.66$, $F(1,64) = 74.26$. The main effect of disturbance was due to a slight tendency for children with disturbance to rate messages more happy overall ($M = 2.86$) than nondisturbed children ($M = 3.05$), $F(1, 64) = 6.33$. These effects are qualified by the interactions to which they contribute.

The Paralanguage × Discrepancy interaction, F(1, 64) = 647.99 (see Table 2 for means and standard deviations), reflects children's reliance on lexical content to resolve discrepancy. Based on Bray and Maxwell's (1985) computational procedure for estimating multiple eta squared ($\eta^2_{mult} = V/s$, where $V$ is the value of the Pillai-Bartlett Trace and s is the number of canonical correlations in the effect matrix), this interaction accounts for 90% of the variance in children's ratings. Across groups, children rated consistent happy messages as more happy than consistent angry messages, $t(66) = 28.03$, but rated angry lexical content-happy paralanguage messages as more angry than happy lexical content-angry paralanguage messages, $t(66) = 14.54$. Pairing either happy or angry para-language with discrepant lexical content produced a significant shift in ratings toward the affect conveyed lexically.

Planned comparisons of the Paralanguage × Discrepancy × Disturbance interaction, $F(1, 64) = 17.70$, indicated that the difference between ratings of consistent happy and consistent angry messages was greater for children with ($M = 2.99$) than for children without disturbance ($M = 2.49$), $t(66) = 3.01$. And the difference between ratings of angry lexical content–happy paralanguage messages and happy lexical content-angry paralanguage messages was significantly greater for children with ($M = 1.93$) than for children without disturbance ($M = 1.16$), $t(66) = 3.01$, with ratings in both cases reflecting the affective valence of lexical content. This finding is consistent with Reilly and Muzekari's (1986) report that between 7 and 9 years of age children with emotional disturbance show a stronger tendency to make affective attributions based on lexical cues than do other children (see Table 3).

The tendency for affective ratings, on the average, to be influenced more by lexical than by paralinguistic content was clear in cases of discrepancy. However, this conclusion is best informed by a consideration of data at the level of individual items and individual children. A lexical bias effect may be sustained primarily by a subset of stimulus items containing ambiguous paralanguage, or it may be an artifact of averaging across individual children employing qualitatively different message resolution strategies. Individual data on children's strategies have not been reported previously (Bugental et al., 1970; McCluskey & Albas, 1981; Reilly & Muzekari, 1979, 1986; Solomon & Ali, 1972).

**Item level performance—**To determine whether lexical bias is exhibited primarily when paralinguistic content is ambiguous, stimuli were divided into ambiguous and unambiguous sets based on children's ratings in Experiment 2. Items for which child interrater agreement was at chance or below were considered ambiguous, whereas items for which child interrater agreement was significantly above chance were considered unambiguous. Lexical content was unambiguous (based on children's ratings in Experiment 1) for all stimulus items. Children's mean ratings, in Experiment 3, are presented as a function of disturbance and paralinguistic ambiguity in Table 4.

For both ambiguous and unambiguous items, children rated angry lexical content-happy paralanguage messages as more angry than happy lexical content–angry paralanguage messages ($Ms = 3.52$ and 2.71), $t(66) = 4.24$ for unambiguous items, and ($Ms = 4.38$ and 1.84), $t(66) = 19.94$ for ambiguous items. This difference was greater for ambiguous ($M = 2.54$) than for unambiguous items ($M = 0.81$), $t(66) = 9.60$, and children with disturbance evinced a greater lexical bias for ambiguous items ($M = 2.96$) than did the nondisturbed sample ($M = 2.13$), $t(66) = 3.53$. This group difference was nonsignificant for unambiguous items.

Across levels of paralinguistic ambiguity, children's affective interpretations most accurately reflect the lexical content of discrepant stimuli. Not surprisingly, this effect is greater when paralinguistic content is ambiguous. Still, with only a single exception, children rated stimuli that contained happy lexical content on the happy end of the scale (below 3) and stimuli that contained angry lexical content on the angry end of the scale (above 3). Next we consider whether lexical bias is an artifact of averaging across individual children employing different response strategies.

**Individual performance—**Examination of Tables 2 and 3 suggests that children's judgments reflect a weighted-averaging strategy favoring lexical content. Rank ordering the stimuli in terms of group ratings supports this notion: Consistent angry messages are given the most angry interpretations followed, in turn, by angry lexical content-happy paralanguage messages, happy lexical content–angry paralanguage messages, and consistent happy messages.

The proportion of children in each group for which this pattern was observed was calculated. In both samples, 52% of children showed the same ranking pattern observed in the summary data, suggesting a weighted-averaging strategy favoring lexical content at the level of individual data. However, this leaves a large proportion of participants unaccounted for in each group. Analysis of the remaining participants revealed that, in many cases, they violated the ranking pattern by placing even greater weight on the lexical content of discrepant messages. For example, in some cases, happy lexical content–angry paralanguage messages were rated as slightly more happy than consistent happy messages or angry lexical content-happy paralanguage were rated as slightly more angry than consistent angry messages. If children employing these strategies are included among those placing greater weight on lexical content, the proportion of children accounted for rises to 82% of nonreferred children and 97% of children with emotional disturbance. The remaining children appeared to use a weighted-averaging strategy favoring paralanguage. A chi-square test of independence indicated that the difference in the distribution of strategies between groups was nonsignificant. These data indicate that children, independent of psychological referral status, are biased toward lexical interpretations of affective discrepancies. This pattern was upheld at the level of both group and individual data.

## EXPERIMENT 4

In Experiment 4, we assessed developmental changes in children's weighting of lexical and paralinguistic information from age 4 to age 10. An important focus was the generalization of effects obtained in Experiment 3 to a new sample of children using a new set of experimental stimuli. The data presented in Experiment 3 suggest that children between 7 and 10 years of age utilize a weighted-averaging strategy that favors lexical content to resolve discrepant communications. In contrast, Dews et al. (1996) reported that children comprehend a speaker's intended meaning in commonly occurring discrepancies such as sarcasm (ironic criticism) and joking (ironic compliments) sometime between 5 and 6 years of age.

However, we expected to observe a weighted-averaging strategy favoring lexical content until at least 7 years of age for two reasons. First, comprehension in the Dews et al. (1996) study was assessed by asking children to choose which meaning best fit the message, the exact lexical content they had just heard or lexical content that was opposite in meaning. For example, when a story character said, "Helpful, isn't he?" the experimenter would ask whether the character meant the person was "helpful and nice" or "selfish and not cooperating" (p. 3073). Children demonstrated their comprehension of discrepant messages by selecting the second, nonliteral, choice. To answer this question correctly, one can either (a) recognize that the intended meaning is the *opposite* of the literal meaning or (b) note that the intended meaning is *not the same* as its lexical content. In principle, a weighted-averaging strategy favoring lexical content could produce this second sort of outcome because this strategy yields different attributions for different combinations of lexical and para-linguistic content (see Tables 2 and 3). Second, the data in Experiment 3 and in earlier research on children's interpretations of discrepancy (Reilly & Muzekari, 1979,1986; Solomon & Ali, 1972; Winner et al., 1987) suggest that a strategy favoring paralanguage is fairly late developing. To assess developmental changes in children's strategies for resolving discrepancy, a subset of data from a larger study of children's understanding of vocal paralanguage was analyzed (Friend, in press).

## Method

**Participants—**Twenty-seven girls were recruited from university housing and from local schools and day care centers. There were 9 girls in each of three age groups: 4-year-olds (*M* = 4;7, range = 4;3 to 4;11); 7-year-olds (*M* = 7;6, range = 7;1 to 7;10); and 10-year-olds (*M* = 10;9, range = 10;2 to 11;4). All children had normal hearing as assessed by parental report. Each child was tested individually in a single-walled sound attenuating chamber in the laboratory.

**Stimuli—**Six new stimulus sentences (3 happy and 3 angry) rated at an interrater agreement of .75 by adults in a previous experiment (Bugental et al., 1970) constituted the lexical content of the stimuli for Experiment 4. Each sentence was read in both happy and angry paralanguage. The experimenter, a female native speaker of American English, produced the stimuli. The stimuli were read into a Shure Model SM81 condenser microphone and recorded using a Teac Model A-3300SX laboratory tape recorder. The microphone was affixed to the experimenter's head at a constant distance of 15.3 cm. Following the reading of each sentence, a reiterant stimulus was generated. The vocal affect of the reiterant stimuli was validated by adults in a previous experiment at a minimum interrater agreement of .75 (Friend & Farrar, 1994).

**Procedure—**Children heard stimuli presented in three counter-balanced, within-subjects conditions. The present analysis was focused on only that subset of children who heard consistent and discrepant speech stimuli first, since our purpose is to clarify developmental changes in interpretations of affective discrepancy. The stimuli were 6 consistent (3 happy lexical content-happy paralanguage, 3 angry lexical content-angry para-language) and 6 discrepant (3 happy lexical content-angry paralanguage, 3 angry lexical content-happy paralanguage) utterances. The stimuli were arranged in both random and reversed order with the restriction that no more than two stimuli representing the same vocal affect (happy or angry) occurred in succession. Each child heard both arrangements (random and reversed), and the order of the arrangements was counterbalanced across children. In addition, reiterant speech versions of the stimuli were included as a manipulation check to ensure that children understood the instructions. Children were told, "On this tape you will hear recordings of a mother talking to her little girl in a toy store. Sometimes she feels happy and sometimes she feels mad. Your job is to tell me when she feels happy and when she feels mad by pointing to one of these faces. Sometimes it will be difficult to understand what she's saying and that's okay. The words are not really important. What's really important is how she sounds." After each stimulus, children were asked to tell the experimenter whether the woman's voice sounded "happy" or "mad" by pointing to one of two schematic faces.

## Results

Performance on the reiterant stimuli revealed that even the youngest children comprehended the task instructions. Data were analyzed in a Paralanguage (2) × Discrepancy (2) × Age (3) mixed models repeated-measures MANOVA. As in Experiment 3, a = .05 and a Bonferroni correction was applied for family-wise comparisons. The analysis revealed a main effect of paralanguage, $F(1, 24) = 106.28$, and several significant interactions: Paralanguage × Age, $F(2, 24) = 7.25$; Discrepancy × Age, $F(2, 24) = 13.99$; Paralanguage × Discrepancy, $F(1, 24) = 102.19$; and Paralanguage × Discrepancy × Age, $F(2, 24) = 5.21$. The Paralanguage × Discrepancy × Age interaction subsumes the lower order interactions and forms the focus of this section (see Table 5 for means and standard deviations).

The interaction was due to a reduction in lexical bias among 10-year-olds relative to younger children. Specifically, 10-year-olds rated happy lexical content-angry paralanguage messages as more angry than did 4-year-olds, $t(16) = 4.29$. Ten-year-olds also rated angry lexical content–happy paralanguage messages as more happy than did 7-year-olds, $t(16) = 3.05$. That is, 10-year-olds' ratings of discrepant stimuli more closely reflected the affect conveyed by vocal paralanguage, whereas 4- and 7-year-olds' ratings more closely reflected the affect conveyed lexically. This trend is apparent across types of discrepancy and in each comparison between 4-and 10-year-olds and 7- and 10-year-olds, although the effect was significant only in those cases listed earlier due to the correction for family-wise error (see Table 5). There were no comparable developmental changes between ages 4 and 7. Table 5 reveals the same weighted-averaging strategy for 4- and 7-year-olds as was observed for 7- to 10-year-olds in Experiment 3. In the current experiment, however, a developmental departure from this pattern is observed for 10-year-olds. Ten-year-olds also appear to utilize a weighted-averaging strategy, but one that favors paralanguage. Data on children's interpretations of individual stimulus items as a function of age appear in Table 6.

**Individual performance—**As in Experiment 3, an examination of the group data for Experiment 4 reveals a weighted-averaging strategy for the interpretation of discrepant messages that appears to favor lexical content over paralinguistic content. Consistent angry messages are given the most angry interpretations followed, in turn, by angry lexical content-happy paralanguage messages, happy lexical content–angry paralanguage messages, and consistent happy messages. The proportion of children in each age group for which this ranking pattern was preserved was calculated. For the 4-year-olds, 6 of the 9 children (67%) followed a weighted-averaging strategy favoring lexical content, and 2 children utilized a strategy favoring paralanguage. For the 7-year-old sample, 5 of 9 children (55%) followed this pattern and 3 children utilized a strategy favoring paralanguage. Finally, in the 10-year-old sample, only one child (11%) followed a strategy favoring lexical content with the majority of 10-year-olds utilizing a weighting strategy that favors paralanguage. With age, a decrease in the primacy of lexical content for resolving affective discrepancies is observed at the level of both group and individual data.

## GENERAL DISCUSSION

The relative salience of lexical and paralinguistic cues to affect in childhood was assessed. In Experiments 1 and 2, we validated stimuli using adult raters and assessed the accuracy and interrater agreement of child judges on lexical and paralinguistic cues presented in isolation. Experiment 3 contrasted competing hypotheses regarding the ubiquitous-ness of a lexical bias in children’s affective judgments, and in Experiment 4 developmental changes were assessed in children’s strategies for resolving discrepancy.

In Experiment 1, the lexical content of stimuli presented in Experiment 3 were validated using adult judges and children’s accuracy and interrater agreement was assessed. Although accuracy was greater than chance for all of the items that children rated in Experiment 1, criterion agreement was met on a subset (79%) of those items rated at criterion by adults. This finding suggests that children make affective judgments based on lexical cues differently than adults, but that sufficient agreement exists to utilize child ratings for stimulus validation in future research.

In Experiment 2, adults validated vocal paralanguage and, again, children’s accuracy and interrater agreement was assessed. Reiterant speech was used to isolate paralinguistic cues by eliminating meaningful lexical content. Children’s judgment accuracy was above chance for a majority of the reiterant stimuli, and the percentage of items on which accuracy was above chance was consistent with the literature on children’s vocal affect detection (Baltaxe, 1991; Dimitrovsky, 1964; Matsumoto & Kishimoto, 1983). However, children and adults were not comparable judges of speaker affect, and children’s interrater agreement was lower for paralinguistic than for lexical cues. Children are accurate judges, on the average, of both lexical and paralinguistic cues, but there is greater variability in their judgments of vocal paralanguage relative to lexical content. Two conclusions, one methodological and one conceptual, follow from these results. First, children’s judgments of vocal paralanguage may be too variable to permit the use of child raters in validation studies. Second, the disparity in child agreement for lexical and paralinguistic cues suggests one possible explanation for the lexical bias effect: that children are simply relatively less skilled in interpreting emotion conveyed paralinguistically. However, because children may find rating voices without words somewhat unusual, we must also consider children’s performance in Experiment 3 in which lexical and paralinguistic content were presented in combination.

Experiment 3 included a comparison sample of children who are frequent recipients of discrepant communication (Blotcky et al., 1982; Bugental et al., 1971). Both discrepancy and disturbance influenced children's interpretations: Both groups placed greater weight on lexical than on paralinguistic content in their interpretations of discrepant messages, and this effect was slightly greater for children with disturbance. The reliance on lexical content in this age range is consistent with the existing literature (Bugental et al., 1970; Friend, in press; Reilly & Muzekari, 1979, 1986; Solomon & Ali, 1972). However, two aspects of children's reliance on lexical content have not been addressed in previous research: the extent to which this reliance remains even when paralanguage is unambiguous and the extent to which it is observed at the level of individual performance. Findings of a lexical bias only in cases of paralinguistic ambiguity or for only a small proportion of children would place important limits on the implications of lexical bias for child development.

In the assessment of paralinguistic ambiguity, we observed that the paralinguistic content of all messages heard by children was unambiguous to adults, although not necessarily unambiguous to children. Children may bring a lower sensitivity to vocal paralanguage, relative to adults, to the task of resolving discrepancy. However, this lower sensitivity cannot fully account for children's lexical bias: Children gave greater weight to lexical than to paralinguistic content even for messages whose paralanguage was rated unambiguously by other children. This tendency was greater when paralanguage was ambiguous. Ambiguity appears to exacerbate the existing developmental bias to weight lexical content more heavily than paralanguage. An early study of the role of lexical ambiguity in adults' interpretations of discrepancy supports this assertion (Argyle et al., 1971). In general, adults'judgments reflected the information conveyed paralinguistically for both lexically ambiguous and lexically unambiguous items, although the effect was greater for ambiguous items. Ambiguity of an ancillary source of affective information appears to enhance existing biases toward basing decisions on a principal source (i.e., lexical content for children and paralinguistic content for adults). The present research was not designed to assess ambiguity systematically, however, and further experiments that address this issue directly are required.

In the present research, a lexical bias occurred even when paralanguage was unambiguous and it was characteristic of the response strategies of individual children. The majority of children used a weighted-averaging strategy favoring lexical information that was also observed at the level of group data. Conceptually, this suggests that lexical bias is the product of a weighting process rather than of exclusive attention to lexical content. That is, children appear to process both cues but attach greater weight to lexical cues. The presence of a lexical bias across items, levels of paralinguistic ambiguity, disturbance, and individual discrepancy resolution strategies suggests that this is a robust developmental phenomenon.

In Experiment 4, we assessed developmental changes in children's strategies for weighting lexical and paralinguistic information. The tendency to weight lexical content more heavily than paralanguage was replicated demonstrating generalizability to a new stimulus set and to a new sample. Both group and individual data in Experiment 4 revealed a developmental progression from a weighted-averaging strategy favoring lexical content at 4 and 7 years of age to one favoring vocal paralanguage at 10 years of age.

This phenomenon has received very limited attention in the empirical literature with a consequent paucity of theoretical approaches to address it. Early accounts assumed, implicitly, that experience with complex messages promotes more adultlike performance (Bugental et al., 1970; Solomon & Ali, 1972). As children are increasingly exposed to complex messages (e.g., sarcasm and joking), they may begin, indeed, to sort out the relevant cues for interpreting them. However, there are several limitations to this approach. First, experience as an explanation presupposes underlying mechanisms that are as yet unspecified. Second, the generalization of lexical bias in Experiment 3 suggests that experience alone does not promote greater weighting of vocal paralanguage. Most important, this explanation does not specify the mechanisms that might account for the emergence of a lexical bias in early childhood.

We speculatively posit two mechanisms that, in conjunction, might produce this effect. The first mechanism concerns children's developing capacity for mental representation. Preschoolers have considerable difficulty on laboratory tasks that require them to reflect on conflicting representations (Flavell, 1988; Friend & Davis, 1993; Gopnik & Astington, 1988). Some accounts of this phenomenon emphasize an information-processing approach. For example, Davis and Pratt (1995) report a positive relation between backward digit span (a measure of central executive capacity) and the understanding of false beliefs. Surian (1995) reports that processing load predicts performance on tasks involving the evaluation and production of ambiguous messages. Frye, Zelazo, and Palfai (1995) suggest that certain complex forms of reasoning are shared by tasks involving conflicting representations of mental states and tasks involving conflicting representations of physical objects. Capacity- or reasoning-based limitations on multiple representations would almost certainly lead younger children to be biased to attend primarily to one source of information (lexical or paralinguistic) in cases of discrepancy.

Flavell suggested that further developments in the understanding of mental representation take place beyond the preschool period, "In short, we are 'at risk'for unconscious egocentrism all of our lives" (Flavell, 1985 as cited in Flavell, 1988, p. 261). Because much of the focus in the literature is on assessing early competence, our knowledge of the types of tasks in which children continue to make representational errors is limited. Judging a speaker's emotion in discrepant communication may constitute just such a task. Other work on children's understanding of affective conflict suggests that the ability to reflect on multiple affective representations is fairly late developing. Harter and Buddin (1987) found that reports of the simultaneous presence of conflicting emotions toward a single target were rare before approximately 11 years of age. Gnepp, McKee, and Domanic (1987) found that 8-year-olds spontaneously reported the possibility of a mixed emotional response only about 50% of the time. A constraint on multiple representations can account for performance on a wide variety of tasks that require the manipulation of conflicting cues. In the current context, this type of constraint might be operationalized as a limited ability to reflect on lexical and paralinguistic cues simultaneously or to shift attention between these representations. However, it cannot account for the directional nature of the observed effect.

A second mechanism is required to explain the consistency with which children attach greater weight to lexical than to paralinguistic content. The posited mechanism concerns the fact that some acoustic variables which contribute to vocal paralanguage (e.g., $F_0$, intensity, and duration) also play a role in facilitating linguistic competence by drawing attention to important lexical items (Fernald & Mazzie, 1991; Garnica, 1977). In fact, Peters (1985) argues that rhythm, stress, and intonation guide the segmentation and extraction of lexical information for an extended period of the language acquisition process. To the extent that children's attention to lexical content is recruited by paralanguage, there may be a developmental bias toward making attributions on the basis of lexical content because it is the subject of focus. This is, in essence, an argument for the processing priority of lexical, relative to paralinguistic, information to accomplish the developmental requirement of proficiency with spoken language.

But how can we account for the judgments of children with disturbance relative to their matched peers? The effect of disturbance was quite small relative to the other effects reported in this research, and individual response strategies did not differ across groups. It appears that the judgments of children with disturbance (across both consistent and discrepant messages) are simply more extreme than the judgments of children without disturbance although in both cases judgments are the product of similar weighting strategies. A parsimonious account of this tendency is that it reflects a response bias toward more extreme affective attributions. This type of account is consistent with a social information processing view of social competence (Crick & Dodge, 1994; Dodge, 1986). In general, however, the more global mechanisms that we have suggested could account for the developmental bias toward lexically based judgments independent of affective disturbance.

The present research provides a methodological clarification and theoretical extension of previous work on children's understanding of affective discrepancy. Children's bias toward lexically based interpretations is noteworthy given the importance of vocal paralanguage in animal communication, mother-infant interaction, and presumably, in human evolution. We have proposed two general mechanisms (one emerging out of the language acquisition process and one reflecting changes in attention and representation) that are implicated in children's lexical bias and that constitute directions for future research.

## Acknowledgments

Data collection for Experiments 1–3 was conducted at the University of South Florida and was supported, in part, by Florida Educational Research and Development Council Award HEXH85. Data collection for Experiment 4 was conducted under the joint sponsorship of the University of California, Irvine and the University of Florida and was supported by funds from the following research grants: NICHD T32-HD07318, NICHD HD-07893, and NIMH R03MH44032. Portions of this research have been presented at the meetings of the Society for Research in Child Development, April 1987 and 1995. We gratefully acknowledge James J. Jenkins, Winifred Strange, Linda Polka, Bruce G. Berg, and Joseph J. Campos for providing facilities for the conduct of these studies and for manuscript preparation. In addition, we are grateful to Paul Ekman, Donna Mumme, Magdalena Smoczynska, and Penelope Vinden for their comments on an earlier version of this manuscript.

## References

Argyle M, Alkema F, Gilmour R. The communication of friendly and hostile attitudes by verbal and non-verbal signals. European Journal of Social Psychology. 1971; 1:385–402.

Baltaxe CA. Vocal communication of affect and its perception in three-to four-year-old children. Perceptual and Motor Skills. 1991; 72:1187–1202. [PubMed: 1961667]

Bateson G, Jackson D, Haley J, Weakland J. Toward a theory of schizophrenia. Behavioral Science. 1956; 1:251–264.

Blotcky AD, Tittler BI, Friedman S. The double-bind situation in families of disturbed children. journal of Genetic Psychology. 1982; 141:129–142. [PubMed: 7142981]

Bray, JH., Maxwell, SE. Multivariate analysis of variance. London: Sage; 1985.

Buck R. Nonverbal communication of affect in children. Journal of Personality and Social Psychology. 1975; 31:644–653. [PubMed: 1159612]

Bugental DE, Kaswan JW, Love LR. Perception of contradictory meanings conveyed by verbal and nonverbal channels. Journal of Personality and Social Psychology. 1970; 16:647–655. [PubMed: 5489500]

Bugental DE, Love LR, Kaswan JW, April C. Verbal-nonverbal conflict in parental messages to normal and disturbed children. Journal of Abnormal Psychology. 1971; 77:6–10. [PubMed: 5100129]

Capelli CA, Nakagawa N, Madden CM. How children understand sarcasm: The role of context and intonation. Child Development. 1990; 67:1824–1841.

Carswell L, White WF. Problems of reporting socioeconomic bias among reading scores and standardized reading tests. Perceptual and Motor Skills. 1984; 58:181–182.

Crick NR, Dodge KA. A review and reformulation of social information processing mechanisms in children's social adjustment. Psychological Bulletin. 1994; 775:74–101.

Darwin, C. The expression of emotion in man and animals. New York: Appleton; 1965/1873.

Davis H, Pratt C. The development of children's theory of mind: The working memory explanation. Australian Journal of Psychology. 1995; 47:25–31.

Dews S, Winner E, Kaplan J, Rosenblatt E, Hunt M, Lim K, Mcgovern A, Qualter A, Smarsh B. Children's understanding of the meaning and functions of verbal irony. Child Development. 1996; 67:3071–3085. [PubMed: 9071771]

Dimitrovsky, L. The ability to identify the emotional meaning of vocal expressions at successive age levels. In: Davitz, JR., editor. The communication of emotional meaning. New York: McGraw-Hill; 1964.

Dodge, KA. A social information processing model of social competence in children. In: Perlmutter, M., editor. The Minnesota Symposium on Child Psychology: Vol. 18. Cognitive perspectives on children's social and behavioral development. Hillsdale, NJ: Erlbaum; 1986.

Fernald A, Mazzie C. Prosody and focus in speech to infants and adults. Developmental Psychology. 1991; 27:209–221.

Flavell, JH. The development of children's knowledge about the mind: From cognitive connections to mental representations. In: Astington, JW.Harris, PL., Olson, DR., editors. Developing theories of mind. New York: Cambridge University Press; 1988.

Friend M. Developmental changes in sensitivity to vocal paralanguage. Developmental Science. in press.

Friend M, Davis T. The appearance-reality distinction: Children's understanding of the physical and affective domains. Developmental Psychology. 1993; 29:907–914.

Friend M, Farrar MJ. A comparison of the validity of content-masking procedures for obtaining judgments of discrete affective states. Journal of the Acoustical Society of America. 1994; 96:1283–1290. [PubMed: 7962995]

Frye D, Zelazo PD, Palfai T. Theory of mind and rule-based reasoning. Cognitive Development. 1995; 10:483–527.

Garnica, OK. Some prosodic and paralinguistic features of speech to young children. In: Snow, CE., Ferguson, CA., editors. Talking to children. Cambridge, England: Cambridge University Press; 1977.

Gnepp J, Mckee E, Domanic JA. Children's use of situational information to infer emotion: Understanding of emotionally equivocal situations. Developmental Psychology. 1987; 23:114–123.

Gopnik A, Astington JW. Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. Child Development. 1988; 59:26–37. [PubMed: 3342716]

Harter S, Buddin BJ. Children's understanding of the simultaneity of two emotions: A five-stage developmental sequence. Developmental Psychology. 1987; 23:388–399.

Ladd DR, Silverman Kea, Tolkmitt F, Bergman G, Scherer KR. Evidence for the independent function of contour type, voice quality, and $F_0$ range in signaling speaker affect. Journal of the Acoustical Society of America. 1985; 78:435–444.

Liberman MY, Streeter LA. Use of nonsense-syllable mimicry in the study of prosodic phenomena. Journal of the Acoustical Society of America. 1978; 63:231–233.

Marcon RA. Socioeconomic versus academic emphasis: Impact on kindergartners' development and achievement. Early Child Development and Care. 1993; 96:81–91.

Matsumoto D, Kishimoto H. Developmental characteristics in judgments of emotion from nonverbal cues. International Journal of Intercultural Relations. 1983; 7:415–424.

Mccluskey KW, Albas DC. Perception of contradictory speech by normal and disturbed boys at various age levels. Journal of Abnormal Psychology. 1981; 90:490–493. [PubMed: 7298998]

Mehrabian A, Wiener M. Decoding of inconsistent communications. Journal of Personality and Social Psychology. 1967; 6:109–114. [PubMed: 6032751]

Mumme DL, Fernald A, Herrera C. Infants' responses to facial and vocal emotional signals in a social referencing paradigm. Child Development. 1996; 67:3219–3237. [PubMed: 9071778]

Nakatani LH, Schaffer JA. Hearing "words" without words: Prosodic cues for word perception. Journal of the Acoustical Society of America. 1978; 63:234–245. [PubMed: 632416]

Peters, AM. Language segmentation: Operating principles for the perception and analysis of language. In: Slobin, DI., editor. The crosslinguistic study of language acquisition: Vol. 2. Theoretical issues. Hillsdale, NJ: Erlbaum; 1985.

Reilly SS, Muzekari LH. Responses of normal and disturbed adults and children to mixed messages. Journal of Abnormal Psychology. 1979; 88:203–208. [PubMed: 447903]

Reilly SS, Muzekari LH. Effects of emotional illness and age upon the resolution of discrepant messages. Perceptual and Motor Skills. 1986; 62:823–829. [PubMed: 3725518]

Rogers PL, Scherer KR, Rosenthal R. Content-filtering human speech: A simple electronic system. Behavioral Research Methods and Instrumentation. 1971; 3:16–18.

Sattler, JM. Assessment of children's intelligence. Philadelphia: W. B. Saunders; 1982.

Scherer, KR. Emotion expression in speech and music. In: Sundberg, J.Nord, L., Carlson, R., editors. Music, language, speech, and brain. London: Macmillan; 1991.

Solomon D, Ali FA. Age trends in the perception of verbal rein-forcers. Developmental Psychology. 1972; 7:238–243.

Surian L. Children's ambiguous utterances: A re-examination of processing limitations on production. Journal of Child Language. 1995; 22:151–169. [PubMed: 7759576]

Svejda M. The development of infant sensitivity to affective messages in the mother's voice (Doctoral dissertation, University of Denver, 1981). Dissertation Abstracts International. 1982; 42:4623B.

Winner E, Windmeuller G, Rosenblatt E, Bosco L, Best E, Gardner H. Making sense of literal and nonliteral falsehood. Metaphor and Symbolic Activity. 1987; 2:13–32.

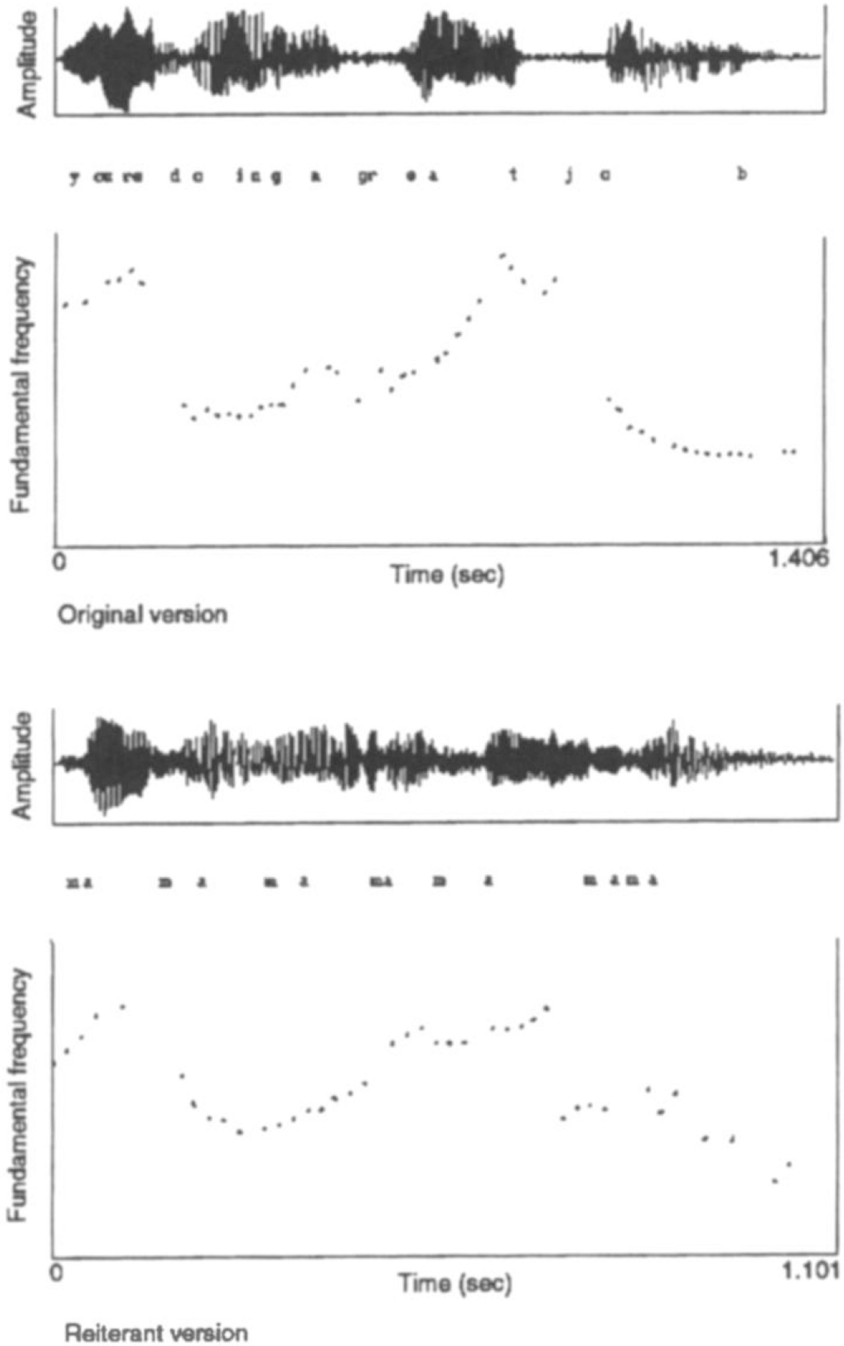


**Figure 1.**
Amplitude waveforms and fundamental frequency contours for original and reiterant versions of the utterance, “You’re doing a great job,” in happy paralanguage.

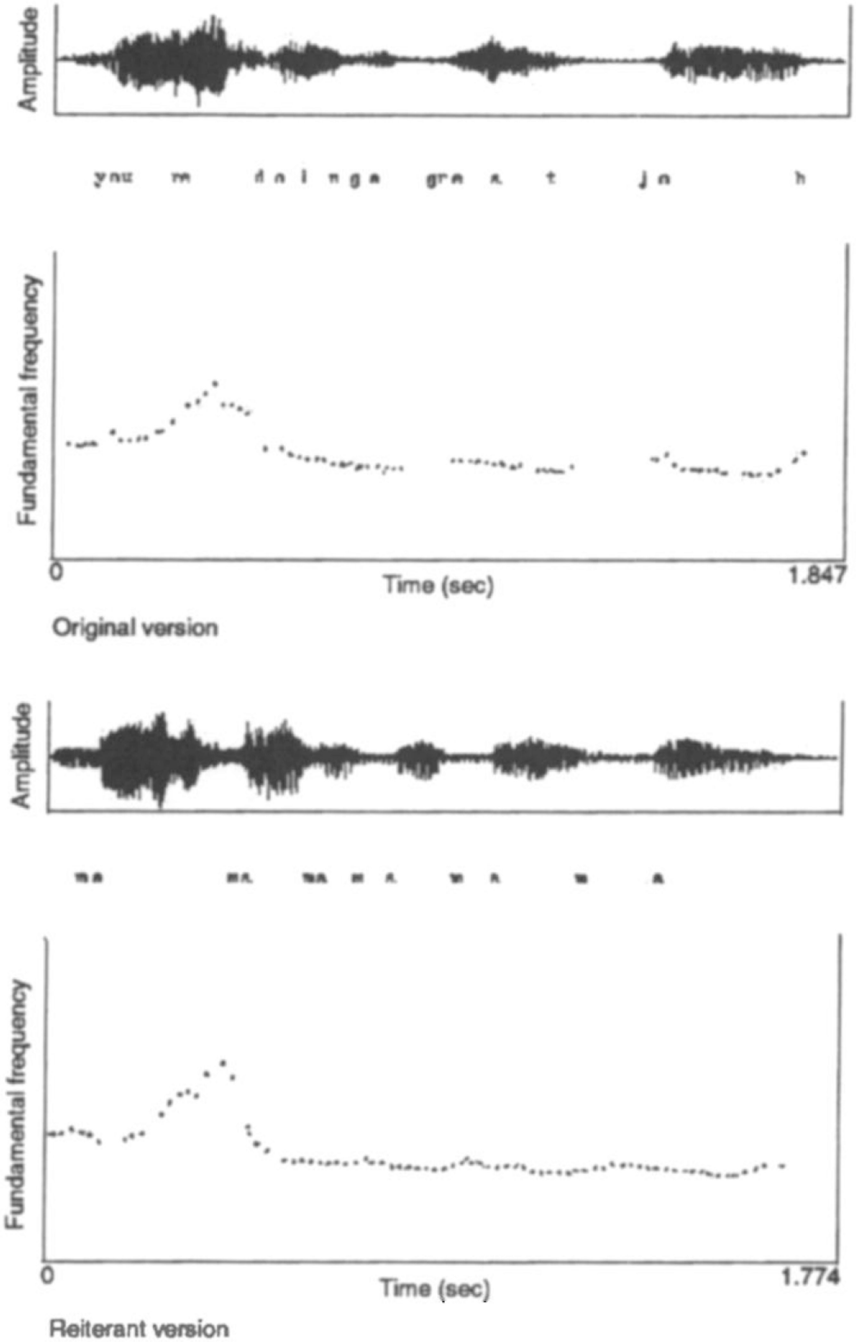


**Figure 2.**
Amplitude waveforms and fundamental frequency contours for original and reiterant versions of the utterance, “You’re doing a great job,” in angry paralanguage.

**Table 1**

Mean Independent Ratings (and Standard Deviations) of Lexical and Paralinguistic Content

| **Stimulus type** | **Lexical ratings** | **Paralinguistic ratings** |
| --- | --- | --- |
| Adults | | |
| Happy lexical content–happy paralanguage | | 1.83 (0.47) |
| | 1.67 (0.39) | |
| Happy lexical content–angry paralanguage | | 4.40(0.16) |
| Angry lexical content–happy paralanguage | | 1.83 (0.47) |
| | 4.08 (0.34) | |
| Angry lexical content-angry paralanguage | | 4.08 (0.23) |
| Children | | |
| Happy lexical content-happy paralanguage | | 2.13(0.77)$_{a}$ |
| | 1.38 (0.28)$_{ac}$ | |
| Happy lexical content-angry paralanguage | | 3.32 (0.56)$_{b}$ |
| Angry lexical content-happy paralanguage | | 2.02 (0.64)$_{c}$ |
| | 4.25 (0.42)$_{b}$ | |
| Angry lexical content-angry paralanguage | | 3.77 (0.69) |

*Note:* Values represent intensity on a 5-point scale (1 = *really happy* and 5 = *really mad*). Items with the same subscript differ at $fw\,\alpha = .05$.

**Table 2**

Mean Affective Ratings (and Standard Deviations) as a Function of Stimulus Type

| | |
|---|---|
| Happy lexical-happy paralanguage | $1.58_a$ (0.44) |
| Happy lexical-angry paralanguage | $2.19_b$ (0.73) |
| Angry lexical-happy paralanguage | $3.74_b$ (0.65) |
| Angry lexical-angry paralanguage | $4.32_a$ (0.45) |

*Note:* A rating of 1 = *really happy* and 5 = *really mad*. Entries with the same subscript differ at $fw\alpha = .05$.

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3**

Mean Affective Ratings (and Standard Deviations) as a Function of Stimulus Type and Disturbance

| | **Disturbance** | |
|---|---|---|
| | **With** | **Without** |
| Happy lexical–happy paralanguage | 1.38 (0.36) | 1.79(0.41) |
| Happy lexical–angry paralanguage | 1.89 (0.58) | 2.48 (0.76) |
| Angry lexical–happy paralanguage | 3.83 (0.68) | 3.65 (0.62) |
| Angry lexical–angry paralanguage | 4.37 (0.44) | 4.28 (0.46) |

*Note:* A rating of 1 = *really happy* and a rating of 5 = *really mad*.

Author Manuscript

Author Manuscript

Author Manuscript

## Table 4

Mean Ratings (and Standard Deviations) of Discrepant Items in Experiment 3 as a Function of Disturbance and Paralinguistic Ambiguity

| | | Disturbance | |
|---|---|---|---|
| **Lexical content** | **Paralanguage** | **With** | **Without** |
| Ambiguous items | | | |
| You’re doing a great job | angry | 1.74(1.21) | 2.12(0.97) |
| I think that’s a real good idea | angry | 1.59 (0.86) | 2.44 (1.02) |
| Oh look this is beautiful | angry | 1.21 (0.54) | 1.94(1.15) |
| Do that once more and you’ll be really sorry | happy | 4.47 (0.82) | 4.29 (0.76) |
| Unambiguous items | | | |
| Great this is just what you’ve been wanting | angry | 2.26(1.54) | 2.68(1.63) |
| You’re growing so fast | angry | 2.67 (1.70) | 3.23(1.42) |
| What makes you think you’ll get away with that? | happy | 3.68(1.07) | 3.12(1.14) |
| You’ll never behave yourself | happy | 3.73(1.21) | 3.67 (0.94) |
| You’ve probably done something wrong again | happy | 3.44(1.05) | 3.50 (0.96) |

*Note:* A rating of 1 = *really happy* and 5 = *really mad*.

Author Manuscript

Author Manuscript

Author Manuscript

**Table 5**

Mean Affective Ratings (and Standard Deviations) in Experiment 4

| | Age (in years) | | |
|---|---|---|---|
| | **4** | **7** | **10** |
| Happy lexical–happy paralanguage | 1.15(0.26) | 1.02 (0.05) | 1.02 (0.05) |
| Happy lexical–angry paralanguage | 1.28 (0.26)$_a$ | 1.48(0.21) | 1.74 (0.19)$_a$ |
| Angry lexical–happy paralanguage | 1.48 (0.33) | 1.67 (0.31 )$_b$ | 1.28 (0.22)$_b$ |
| Angry lexical–angry paralanguage | 1.96(0.11) | 2.00 (0.00) | 1.98 (0.05) |

*Note:* A rating of 1 = *happy* and 2 = *mad*. Entries with the same subscript differ at *fw* $\alpha$ = .05.

**Table 6**

Mean Affective Ratings (and Standard Deviations) of Individual Items in Experiment 4

| | | Age (in years) | | |
|---|---|---|---|---|
| **Lexical content** | **Paralanguage** | **4** | **7** | **10** |
| You'll never behave yourself | angry | 1.89 (0.33) | 2.00 (0.00) | 2.00 (0.00) |
| | happy | 1.39 (0.49) | 1.67 (0.43) | 1.22 (0.36) |
| Don't play around with me child | angry | 2.00 (0.00) | 2.00 (0.00) | 1.94(0.17) |
| | happy | 1.39 (0.42) | 1.50 (0.43) | 1.11 (0.33) |
| You're being punished; you stay right there | angry | 2.00 (0.00) | 2.00 (0.00) | 2.00 (0.00) |
| | happy | 1.67(0.35) | 1.83(0.25) | 1.50(0.35) |
| Oh good you got them all | angry | 1.39 (0.49) | 1.78 (0.36) | 1.94(0.17) |
| | happy | 1.11 (0.33) | 1.06(0.17) | 1.00 (0.00) |
| You play very well | angry | 1.27(0.26) | 1.61 (0.42) | 1.89 (0.22) |
| | happy | 1.17(0.25) | 1.00 (0.00) | 1.05(0.17) |
| You're my favorite person | angry | 1.17(0.25) | 1.06(0.17) | 1.39 (0.42) |
| | happy | 1.17(0.35) | 1.00 (0.00) | 1.00 (0.00) |

*Note:* A rating of 1 = *happy* and 2 = *mad*.