Bilingualism: Language and Cognition 14 (4), 2011, 562–569 © Cambridge University Press 2011 doi:10.1017/S1366728910000404

Bilingualism and children's use of paralinguistic cues to interpret emotion in speech*

W. QUIN YOW
ELLEN M. MARKMAN
Stanford University

(Received: December 10, 2009; final revision received: May 4, 2010; accepted: August 23, 2010; First published online 28 April 2011)

Preschoolers tend to rely on what speakers say rather than how they sound when interpreting a speaker's emotion while adults rely instead on tone of voice. However, children who have a greater need to attend to speakers' communicative requirements, such as bilingual children, may be more adept in using paralinguistic cues (e.g. tone of voice) when interpreting a speaker's affect. We explored whether bilingual children are better able than monolingual children to use paralinguistic cues when interpreting a speaker's emotion. While monolingual and bilingual children were equally capable of identifying emotion using affective information in low-pass filtered speech stimuli (Study 1), bilingual children were better able than monolingual children to use tone of voice when judging emotion in natural speech when content was clear (Study 2).

Keywords: bilingualism, tone of voice, emotion

A given (literal) utterance can be interpreted in many different ways (e.g. Bolinger, 1989; Labov & Fanshel, 1977). Adults often use a wide range of cues, including social and contextual cues, to understand the intended meaning of an utterance, especially a potentially ambiguous one. Adults rely on paralinguistic cues (e.g. tone of voice, pitch, speaking rate) to gain insights into a speaker's communicative intent including whether the speaker intends an utterance to signal request, approval, attention, or irony (Ackerman, 1986; Capelli, Nakagawa & Madden, 1990; Cutler, 1974; de Groot, Kaplan, Rosenblatt, Dews, & Winner, 1995; Egan, 1980; Fernald, 1989; Kreuz, 1996, 2000; Milosky & Ford, 1997).

Paralinguistic cues are an indispensable means by which adults express their thoughts, emotions, and attitudes to one another (Fussell & Moss, 1998; Goldie, 2002; Ortony, 1975; Roberts & Kreuz, 1994). These cues can signal the intentionality of an emotion; for example, they can disclose the speaker's state of mind beyond the bounds of one's literal utterance. Even babies often make use of simple paralinguistic cues, such as line-of-regard and tone of voice, to interpret a speaker's emotion and referential intent (e.g. Baldwin & Moses, 1994). According to Relevance Theory, communication involves the expression and extraction of information that is relevant to a speaker's intended meaning. In situations where there is ambiguity in interpreting a speaker's thoughts and feelings (e.g. irony, sarcasm, pretense, deception), adult listeners expect the speaker to provide disambiguating cues such as intonation and facial expression that the listeners would then use to interpret the utterances (Bryant & Fox Tree, 2002; Sperber & Wilson, 1986). Indeed, past research has shown that adults often turn to such nonverbal cues in ambiguous situations to figure out a speaker's underlying intent (Cutler, 1974; Ekman, 1985; Kreuz & Roberts, 1995; Rockwell, 2000) and emotion (Argyle, Alkema & Gilmour, 1971; Mehrabian & Wiener, 1967; Morton & Trehub, 2001).

In contrast, young children generally find it hard to use paralinguistic cues in ambiguous contexts, such as when paralanguage conflicts with literal meanings. Six-year-olds have difficulty in recognizing sarcasm (Capelli et al., 1990; Demorest, Meyer, Phelps, Gardner & Winner, 1984) and irony (Ackerman, 1982; Andrews, Rosenblatt, Malkus, Gardner & Winner, 1986). When lexical content was pitted directly with paralanguage (e.g. a happy event expressed in a sad voice), children, unlike adults, were shown to give more weight to content than paralanguage (Friend, 2000; Friend & Bryant, 2000; Morton & Munakata, 2002; Morton & Trehub, 2001; Solomon & Ali, 1972). For example, Morton and Trehub (2001) asked children and adults to judge whether a speaker was happy or sad. Participants heard sentences describing happy and sad situations in either a happy or sad voice. When content and paralanguage matched, both children and adults could accurately identify happy and sad sentences. When the cues conflicted, adults overwhelmingly relied on how

* We are especially grateful to Bruce Morton for generously sharing his stimuli. We thank the children and parents who participated and to the teachers and staff of Bing Nursery School. We thank Adrienne Sussman and Suejung Shin for helping to collect and code the data. Portions of this work were previously presented at the Biennial Meeting of the Society of Research in Child Development in Boston (March 2009). This work was partially supported by the Tan Kah Kee Postgraduate Scholarship to the first author.

Address for correspondence:
W. Quin Yow, Department of Psychology, Stanford University, 450 Serra Mall, Jordan Hall Building 420, Stanford, CA 94305-2130, USA
quin@psych.stanford.edu

speakers sounded, while four-year-olds almost exclusively judged speakers' emotion from what they said. Although children are able to use paralanguage to judge emotion when primed or given explicit instructions and feedback to do so (Morton, Trehub & Zelazo, 2003), there is a propensity for children to rely on lexical content over paralanguage to judge emotion when these cues conflict.

Previous studies suggest that children's ability to infer a speaker's communicative intent is a prerequisite for understanding non-literal expressions such as irony and sarcasm (Andrews et al., 1986; Sullivan, Winner & Hopfield, 1995; Winner & Leekam, 1991). A speaker's communicative intent can often be determined by cues such as intonation, especially when a speaker intends a non-literal interpretation (de Groot et al., 1995). It follows, then, that children who have a greater sensitivity to a speaker's communicative intent may be better able to use intonation to interpret non-literal expressions than children who do not. Several studies have documented that children growing up in a bilingual environment have a heightened sensitivity to a speaker's communicative intent compared to monolingual children as they have to frequently monitor the communicative context to determine what language a speaker is using and how to respond appropriately (Ben-Zeev, 1977; Comeau, Genesee & Lapaquette, 2003; Comeau, Genesee & Mendelson, 2007; Cummins & Mulcahy, 1978; Genesee, Tucker & Lambert, 1975; Siegal, Iozzi & Surian, 2009; Yow & Markman, 2007, 2009). Therefore, bilingual children may be better able to use paralinguistic cues to interpret a speaker's emotion than monolingual children in contexts where children typically tend to rely on content over paralanguage to evaluate emotion. When there is no conflict between intonation and content, monolingual children have been shown to successfully use tone of voice to determine a speaker's emotion (Morton & Trehub, 2001). Thus, the bilingual advantage should be expected only when intonation and content conflict.

Study 1 first explores whether there is a difference between monolingual and bilingual children in their ability to identify paralanguage per se when there is no conflict with lexical content. We used Morton and Trehub's (2001) filtered-speech stimuli, where affective information (fundamental frequency and speaking rate) was retained but content was unintelligible. We expected that monolingual and bilingual children would be equally proficient in using paralanguage to identify a speaker's emotion when there is no incompatible content. Study 2 addresses our primary question by adapting Morton and Trehub (2001) and Morton et al.'s (2003) speech stimuli where happy and sad situations were recorded in either happy or sad voices. For conflicting sentences, we predicted that bilingual children would be better able than monolingual children to use paralinguistic cues over content to interpret a speaker's emotion.

Study 1

Method

Participants

Thirty-two English monolingual and bilingual four-year-olds from a preschool in Palo Alto participated in this study. Sixteen were monolinguals (six males; mean age = 4.47; range = 4.04–4.92) and 16 were bilinguals (seven males; mean age = 4.36; range = 4.12–4.94). A language questionnaire was sent to the parents via the school that asked for information about the language first acquired by the child, the language used by the parents and caregivers, and the amount of time (percentage of exposure per week) the child was exposed to each language. Bilingual children were determined as those with at least 30% exposure to one of two languages weekly since birth. The 16 bilingual children in the study were reported to have regular exposure to another language besides English, such as Spanish (n = 6), Mandarin (n = 4), Korean (n = 2), French, Japanese, Italian, and German (n = 1 per language), mainly either from parents or a nanny.

All children were recruited from the same university lab school, lived in Palo Alto and its neighboring areas, and were mostly middle-to-upper class. To verify that monolingual and bilingual children were drawn from the same socioeconomic status (SES) population, we followed the procedure reported by Buck, Small, Schisterman, Lyon and Rogers (2000), Furth, Garg, Neu, Hwang, Flush and Powe (2000), Rathore, Masoudi, Wang, Curtis, Foody, Havranek and Krumholz (2006), Vaillancourt, Lui, Maio, Wells and Stiell (2008), Ward (2008), and Westenberg, Siebelink, Warmenhoven and Treffers (1999), and used participants' residential addresses to obtain an estimated value of each family's dwelling from an internet website that provides real estate information such as home prices and home values (www.zillow.com). Property value is correlated with other measures of socioeconomic status such as median income per household and thus is regarded as a valid proxy for socioeconomic status (e.g. Clarke, Schellenbaum & Rea, 2005; Hallstrom, Boutin, Cobb & Johnson, 1993). Using this method, we calculated the median, mean, and variance property valuation. The ratio of the median property valuation between monolingual and bilingual children was 1:1.20 and Mann-Whitney U-test confirmed that these two groups of children came from the same SES backgrounds, $Z = -.44$, $P = .66$. The ratio of the means was 1:1.01, and t-tests showed no significant differences between these two groups of children based on the estimated property valuations, $t(26) = .028$, $p = .98$. The ratio of the variances was 1:0.84 and the Levene test of equality in variances confirmed that the two group variances of estimated property valuations did not differ significantly from each other, $F(1,26) = .11$, $p = .74$.

Materials

The same filtered speech stimuli from Morton and Trehub (2001) were presented to each child on a Macintosh computer using the program Matlab 7.5 (MathWorks, 2007) with Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). There were two pre-designated buttons on the keyboard for children to record their responses. The speech stimuli were normal spoken utterances describing happy and sad situations in either happy or sad paralanguage but low-pass filtered at 500 Hz (90 dB per octave roll off). However, since such filtering greatly reduced the variability of sentences and made the task somewhat boring for young children, the number of sentences used in Study 1 was reduced to 32 (from 40 sentences used in Morton and Trehub). Half of the sentences were happy and half were sad. Of the 32 sentences, two (one happy, one sad) were used as practice trials and the remaining sentences were used in the test phase. Children were randomly assigned to four different pre-determined randomized orders, counterbalanced for type of sentence (happy vs. sad) as the first practice trial and the first experimental trial, with no more than three happy or sad sentences in a row. Response time, in milliseconds, was collected as the time from which each speech stimulus has ended to the time a computer key has been pressed.

Procedure

The procedure was similar to Morton and Trehub's (2001) study with expanded instructions. Children were told to listen carefully to the experimenter's friend, Marianne, who would be talking about many different things. Sometimes the speaker felt happy and sometimes she felt sad. Children were also told that the speaker was being a bit funny so she sounded different from her normal voice. Children were asked to press one of the two pre-designated buttons after they heard each sentence: "happy" button if they thought the speaker was feeling happy or "sad" button if they thought the speaker was feeling sad. They were reassured that sometimes it would be difficult to understand what the speaker is saying and that is okay. After two practice trials, 30 sentences were then presented in three blocks of ten. Each block consisted of five happy and five sad sentences. There were prompts in between blocks to remind the children that they had to press the happy button if they thought she sounded happy and vice versa.

Results and discussion

We predicted that monolingual and bilingual preschoolers would be equally able to label the affective paralanguage in the filtered utterances. Children were given a score of 0–30 that reflects the total number of times they chose the correct response. An omnibus univariate ANOVA was conducted and revealed no effect of order, gender, or type of affect (all *p*s $>$ .10). There was also no significant correlation between SES and performance in the task, $r = -.090$, $p = .65$.

Table 1. *Mean total number of correct responses (out of 30) and average median reaction times (RT) in Study 1 (Standard Deviations in parentheses).*

Language status	Mean correct responses	Average median RT (s)
Monolingual	17.81 (4.39)	1.98 (1.19)
Bilingual	18.81 (5.60)	1.40 (0.58)

The mean number of correct responses was 18.30 and the average median reaction time (RT) for correct trials was 1.69 s. Median RT is often used in place of mean RT because it is not sensitive to occasional extreme values (e.g. Hays, 1973; Mitchell, Zhou, Chavez & Guzman, 1992). Table 1 presents the mean total number of correct responses and average median RTs by language status.

As predicted, there was no difference between monolingual and bilingual preschoolers' ability to use paralanguage to interpret a speaker's emotion when content was removed from the speech. A univariate ANOVA by language status (monolingual vs. bilingual) was conducted. No significant effect was found, $F(1,30) = .32$, $p > .57$. We also compared performance against chance. One-sample *t*-tests revealed above chance performance across children, monolingual: $t(15) = 2.56$, $p < .05$; bilingual: $t(15) = 2.73$, $p < .05$. Therefore monolingual and bilingual children alike were able to use paralanguage to label a speaker's emotion.

Kolmogorov-Smirnov test (K-S) of normality suggested that not all subgroups of the reaction time data were normally distributed (monolinguals: $K\text{-}S = .23$, $p = .026$; bilinguals: $K\text{-}S = .17$, $p = .20$). Normality was achieved after median RTs were log-transformed ($p = .20$ for both groups). A univariate ANOVA by language status (monolingual vs. bilingual) was conducted on the log-transformed median RTs. No significant effect of language status was found, $F(1,30) = 2.18$, $p > .15$ (monolingual: $M = .52$, $SD = .59$; bilingual: $M = .26$, $SD = .42$). Monolingual and bilingual children did not differ significantly in their response speed.

Our results showed that monolingual and bilingual children are equally capable of identifying emotion based on paralinguistic cues (e.g. fundamental frequency and speaking rate) when no content was discernable. In Study 2, we explored, in a natural speech context, whether bilingual children are better able than monolingual children to use paralinguistic cues to judge emotion when those cues conflict with semantic content.

Study 2

Method

Participants

Thirty-two English monolingual and bilingual four-year-olds from a preschool in Palo Alto participated in this study. Sixteen were monolinguals (eight males; mean age = 4.62; range = 4.11–4.99) and 16 were bilinguals (eight males; mean age = 4.71; range = 4.35–4.99). These children did not participate in Study 1. Language questionnaire and questionnaire coding were the same as used in Study 1. The 16 bilingual children in the study were reported to have regular exposure to another language besides English since birth, such as Spanish (n = 7), Farsi (n = 2), Mandarin, Japanese, Tibetan, Russian, Hungarian, Swiss-German, Korean (n = 1 per language), mainly either from parents or a nanny.

All children were recruited from the same university lab school, lived in Palo Alto and its neighboring areas, and were mostly middle-to-upper class. Using the same method as Study 1, we calculated the mean, median, and variance of property valuation. The ratio of the median property valuation between monolingual and bilingual children was 1:0.65 ($Z = -.71$, $P = .58$), the ratio of the means was 1:0.77 ($t(28) = .99$, $p = .33$), and the ratio of the variances was 1:0.45 ($F(1,28) = .32$, $p = .58$), all of which indicated that both groups of children came from the same SES background.

Materials

The speech stimuli were presented on a Macintosh computer. The same 40 spoken utterances were used as those in Morton and Trehub (2001). There were 10 sentences describing happy situations (e.g. "my mommy give me a treat") and 10 describing sad situations (e.g. "my dog ran away from home"). Each sentence was recorded twice, once with happy paralanguage and once with sad paralanguage. In addition, four utterances with neutral content from Morton et al. (2003) were included as practice trials (e.g. "I live in Mississauga"), two of which were recorded with happy paralanguage and two with sad paralanguage. Children were randomly assigned to eight different pre-determined randomized orders, counterbalanced for type of sentence (happy vs. sad) and condition (consistent vs. discrepant) as the first experimental trial, with no more than three congruent or incongruent trials in a row and no more than three happy or sad sentences in a row. Response time, in milliseconds, was collected as the time from which each speech stimulus has ended to the time a computer key has been pressed.

Procedure

The same procedures as Study 1 were used with the exception that children were not told that the speaker would sound different from her normal voice. In addition, children had four practice trials with neutral sentences. Forty sentences were then presented in four blocks of ten. Each block consisted of five happy and five sad sentences. There were prompts in between blocks to remind children that they had to press the happy button if they thought she sounded happy and vice versa.

Table 2. *Mean total number of correct responses (out of 20) and average median reaction times (RT) in Study 2 (Standard Deviations in parentheses) by condition.*

Condition	Language status	Mean correct responses	Average median RT (s)
Consistent	Monolingual	18.19 (1.60)	1.70 (1.04)
	Bilingual	18.13 (2.42)	1.56 (0.65)
Discrepant	Monolingual	5.39 (3.53)	2.03 (1.10)
	Bilingual	11.75 (4.89)	2.00 (0.67)

Results and discussion

We predicted that while monolingual children can accurately interpret a speaker' emotion when content is consistent with paralanguage (consistent condition), bilingual children would be better able than monolingual children to use paralanguage when content and paralanguage did not match (discrepant condition). For the discrepant condition, we scored a judgment based on paralanguage as correct because that is what adults do (see Morton & Trehub, 2001). Hence, children were given a score of 0–20 for each condition that reflects the total number of times they chose the correct response. An omnibus repeated measures ANOVA was conducted and revealed no effect of order, gender, or type of affect (all *p*s > .10). There was also no significant correlation between SES and performance in either the consistent or discrepant conditions, $r = .19$ and $.23$, $p = .32$ and $.21$, respectively. Table 2 presents the mean number of correct responses and average median RTs for consistent and discrepant conditions.

A 2 (condition: consistent vs. discrepant) × 2 (language status: monolingual vs. bilingual) repeated measures ANOVA was conducted. There was a significant main effect of condition, $F(1,30) = 148.02$, $p < .01$. Children scored significantly higher in the consistent condition than in the discrepant condition (see Table 2). There was also a significant main effect of language status, $F(1,30) = 12.75$, $p < .01$. Bilingual children obtained significantly higher scores than monolingual children. As predicted, these main effects were modulated by a significant interaction between condition and language status, $F(1,30) = 16.62$, $p < .01$. When content and paralanguage matched, all children identified happy and sad sentences equally well, $t(30) = .086$, $p > .93$. But when content conflicted with paralanguage, monolingual children relied on content while bilingual children were more willing to use paralanguage to judge emotion, $t(30) = 4.22$, $p < .001$.

We wanted to examine whether the effects we found could be generalized over items. To do this, we treated items as a random factor with each item receiving a score from 0 to 16 depending on how many monolingual or bilingual children in each condition answered the item correctly. We conducted a 2 (condition: consistent vs. discrepant) × 2 (language status: monolingual vs. bilingual) repeated measures ANOVA on scores for each item. All of the results over items paralleled those found over subjects. There was a significant main effect of condition, $F(1,19) = 120.47$, $p < .001$. Items received higher scores in the consistent condition than discrepant condition ($M = 14.48$, $SE = .33$; $M = 6.75$, $SE = .46$, respectively). There was also a main effect of language status, $F(1,19) = 84.90$, $p < .001$. Items were more often answered correctly when children were bilingual than monolingual ($M = 11.98$, $SE = .26$; $M = 9.25$, $SE = .22$, respectively). Most importantly, there was a significant interaction effect between condition and language status, $F(1,19) = 46.28$, $p < .001$. Paired-sample *t*-tests revealed that when content and paralanguage were consistent with each other, items were answered correctly by both monolingual and bilingual children, $t(19) = .85$, $p > .40$. But when content was in conflict with paralanguage, items received higher scores when children were bilingual than monolingual, $t(19) = -9.02$, $p < .001$. Thus, all of the effects we found in our main subject analyses could be generalized over items.

Turning back to analyses over subjects, we compared the performance of monolingual and bilingual children against chance. One-sample *t*-tests revealed above chance performance in the consistent condition across all children (monolingual: $t(15) = 20.46$, $p < .001$; bilingual: $t(15) = 18.13$, $p < .001$). The pattern was quite different when paralanguage and content were in conflict. Monolingual children were below chance in performance, $t(15) = -5.23$, $p < .001$; they significantly relied on content when judging the speaker's emotion. In contrast, bilingual children were not different from chance, $t(15) = 1.43$, $p > .17$. Bilingual children were more reliant on paralinguistic cues than monolingual children, but were not as consistent as the adults (see Morton & Trehub, 2001).

Following Morton and Trehub (2001), we also grouped children in the discrepant condition into one of three categories depending on how consistently a child focused on content or paralanguage and conducted a chi-square test. The children were grouped into: Content Focus (scores of 0–6), Mixed Focus (scores of 7–13), and Paralinguistic Focus (scores of 14–20). The distribution of data for monolingual children in our study is comparable to that of Morton and Trehub's four-year-old data (see Table 3). However, the distribution of data for bilingual children showed less reliance on content. A chi-square analysis of our Study 2 data confirmed a significant interaction between language status and category, $\chi^2(2, n = 32) = 10.57$, $p < .01$.

Table 3. *Frequency distribution of children's responses to sentences with conflicting cues in Study 2 (% in parentheses).*

	Content Focus (0–6)	Mixed Focus (7–13)	Paralinguistic Focus (14–20)
Monolingual 4-year-olds (n = 16)	12 (75.0%)	3 (18.7%)	1 (6.3%)
Bilingual 4-year-olds (n = 16)	3 (18.7%)	7 (43.8%)	6 (37.5%)
Morton & Trehub's (2001) 4-year-olds (n = 21)	17 (80.9%)	3 (14.3%)	1 (4.8%)

As in Study 1, Kolmogorov-Smirnov test of normality suggested that not all subgroups of the RT data were normally distributed (for consistent trials – monolinguals: $K\text{-}S = .24$, $p = .018$ and bilinguals: $K\text{-}S = .21$, $p = .055$; for discrepant trials – monolinguals: $K\text{-}S = .14$, $p = .20$ and bilinguals: $K\text{-}S = .15$, $p = .20$). Normality was achieved after median RTs were log-transformed (all $ps > .11$). A 2 (condition: consistent vs. discrepant) × 2 (language: monolingual vs. bilingual) repeated measures ANOVA was conducted on the log-transformed median RTs. There was a significant main effect of condition, $F(1,30) = 30.77$, $p < .01$. Children were significantly faster in responding to consistent trials compared to discrepant trials (see Table 2). This is consistent with Morton and Trehub's (2001) findings that children found consistent trials to be easier than discrepant trials. There was no significant effect of language status, $F(1,30) = .033$, $p > .85$. Both monolingual and bilingual children were equally fast to respond. No other significant effects were found.

In sum, as expected, monolingual children relied on content over paralanguage when judging emotion, thus replicating the findings of Morton and Trehub (2001). Bilingual children, on the other hand, showed an early emerging ability to use paralinguistic cues over content but they were not as able as adults to do this consistently.

General discussion

Our results provide evidence that while monolingual and bilingual children are equally capable of using tone of voice to identify emotion when there is no conflicting content, bilingual children are better able than their monolingual peers to judge emotion when content is in conflict with tone of voice. In Study 1, children were presented low-pass filtered speech stimuli removed of any comprehensible content but with affective information retained. Monolingual and bilingual children were equally accurate in identifying emotion based only on tone of voice. In Study 2, children heard natural speech stimuli where content and paralanguage were either consistent or discrepant. When tone of voice and content were consistent, monolingual and bilingual children were equally capable of identifying the emotion. In contrast, bilingual children were more adult-like than monolingual children in using intonation to interpret the emotion of a speaker when it conflicted with lexical content.

Our results replicate previous findings in that children are able to identify emotion using paralinguistic cues in speech when there is no conflict with content. Four-year-olds typically can infer emotion from a range of paralinguistic cues such as pitch, contour, and rate of speaking (Borke, 1971; Dimitrovsky, 1964). At the same time, our findings also correspond to previous research which found that preschoolers are less able to use paralinguistic cues in construing others' emotion when the cues conflict with lexical meaning of the utterances (e.g. Friend, 2000; Friend & Bryant, 2000; Morton & Trehub, 2001). However, most importantly, our findings suggest that bilingual children differ from monolingual children in that they may have a burgeoning ability to use paralinguistic cues to interpret the emotion of a speaker when these cues conflict with lexical content of the utterance.

It is possible that a bilingual advantage in selective attention and control could, in part, explain the present results. According to cognitive complexity and control theory (Frye, Zelazo & Burack, 1998; Zelazo & Frye, 1998), difficulty in relying on tone of voice over content arises because it requires children to keep in mind two possible rules of interpretation, and to formulate and use a higher order rule to inhibit their prepotent tendency to rely on content. And, in fact, several studies have documented improved inhibitory control in bilinguals (e.g. Bialystok, 1999; Bialystok & Martin, 2004; Kovacs & Mehler, 2009; Martin-Rhee & Bialystok, 2008; see Costa, Hernández & Sebastián-Gallés, 2008, for an advantage in bilingual adults). Thus, inhibitory control may help account for the bilingual advantage in our task. However, Morton et al. (2003) have demonstrated that reversing children's bias towards content could be achieved relatively easily without reducing the cognitive control demands of the task. They found that children could be readily primed to judge emotion using paralanguage by first hearing neutral sentences recorded in either happy or sad voices. Thus, their results suggest that cognitive control may not be the key reason young children spontaneously rely on content to interpret the speaker's emotion.

Another explanation for children's reliance on content is that they lack the sensitivity to respond appropriately to the different demands of various listening contexts (Mazzocco, 1999; Solomon & Ali, 1972). This explains why with priming or explicit instructions and feedback to increase awareness of the use of paralinguistic cues, children were able to judge emotion by using paralanguage (Morton et al., 2003). Importantly, we propose that growing up bilingual provides a natural environment for children to learn about the changing communicative demands in a social context. In dealing with the learning and use of two language systems, young bilingual children often have to distinguish the different languages being spoken to and determine the appropriate circumstances to switch between languages in conversations (Hoffman, 1991; Saunders, 1983). They may be more attentive to various communicative cues to help them understand the communicative context and how they should respond, including which language to use with which speaker under what context. Thus, growing up in an environment where one has to regularly monitor the communicative context may heighten children's sensitivity to the use of communicative cues in interpreting speakers' intent. Nevertheless, it remains plausible that both sensitivity to communicative cues (ability to utilize expressions of affect as a disambiguating cue to ambiguous communicative situation) and inhibitory control skills (ability to ignore content and utilize tone of voice) contribute to bilingual children's ability to use paralanguage over lexical content to judge emotion.

In sum, adults encounter non-literal expressions in their everyday communicative contexts where speakers do not always say what they mean. Adults often use a wide range of cues to help them understand what a speaker is trying to communicate. While most adults are able to use paralinguistic cues (e.g. tone of voice) to interpret ambiguous utterances, young children typically make literal interpretations based on what it is said. Our results suggest that bilingual children have an incipient adult-like ability to overcome their reliance on content and use other cues such as tone of voice to interpret a speaker's emotion. Bilingual children's use of paralanguage to interpret emotion, particularly when it conflicts with lexical content, may underscore their heightened sensitivity to a speaker's communicative intent.

References

Ackerman, B. (1982). Contextual integration and utterance interpretation: The ability of children and adults to interpret sarcastic utterances. *Child Development, 53,* 1075–1083.

Ackerman, B. (1986). Children's sensitivity to comprehension failure in interpreting a nonliteral use of an utterance. *Child Development, 57,* 458–497.

Andrews, J., Rosenblatt, E., Malkus, U., Gardner, H., & Winner, E. (1986). Children's abilities to distinguish metaphoric and ironic utterances from mistakes and lies. *Communication and Cognition, 19,* 281–298.

Argyle, M., Alkema, F., & Gilmour, R. (1971). The communication of friendly and hostile attitudes by verbal and nonverbal signals. *European Journal of Social Psychology, 1,* 385–402.

Baldwin, D. A., & Moses, L. J. (1994). Early understanding of referential intent and attentional focus: Evidence from language and emotion. In C. Lewis & P. Mitchell (eds.), *Childrens' early understanding of mind: Origins and development*, pp. 133–156. Hillsdale, NJ: Lawrence Erlbaum.

Ben-Zeev, S. (1977). The effect of bilingualism in children from Spanish–English low economic neighborhoods on cognitive development and cognitive strategy. *Working Papers in Bilingualism, 14,* 83–122.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70,* 636–644.

Bialystok, E., & Martin, M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sorting task. *Developmental Science, 7,* 325–339.

Bolinger, D. W. (1989). *Intonation and its uses*. Palo Alto, CA: Stanford University Press.

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

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Bryant, G. A., & Fox Tree, J. E. (2002). Recognizing verbal irony in spontaneous speech. *Metaphor and Symbol, 17 (2),* 99–117.

Buck, G. M., Small, M. E., Schisterman, E. F., Lyon, N. R., & Rogers, B. T. (2000). Extreme prematurity and school outcomes. *Paediatric and Perinatal Epidemiology, 14 (4),* 324–331.

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

Clarke, S. O., Schellenbaum, G. D., & Rea, T. D. (2005). Socioeconomic status and survival from out-of-hospital cardiac arrest. *Academic Emergency Medicine, 12 (10),* 941–947.

Comeau, L., Genesee, F., & Lapaquette, L. (2003). The modeling hypothesis and child bilingual codemixing. *International Journal of Bilingualism, 7 (2),* 113–126.

Comeau, L., Genesee, F., & Mendelson, M. (2007). Bilingual children's repairs of breakdowns in communication. *Journal of Child Language, 34,* 159–174.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106 (1),* 59–86.

Cummins, J., & Mulcahy, R. (1978). Orientation to language in Ukrainian–English bilingual children. *Child Development, 49,* 479–482.

Cutler, A. (1974). On saying what you mean without meaning what you say. *Papers from the Tenth Regional Meeting, Chicago Linguistic Society*, pp. 117–127. Chicago: Department of Linguistics, University of Chicago.

de Groot, A., Kaplan, J., Rosenblatt, E., Dews, S., & Winner, E. (1995). Understanding versus discriminating nonliteral utterances: Evidence for a dissociation. *Metaphor and Symbolic Activity, 10,* 255–273.

Demorest, A., Meyer, C., Phelps, E., Gardner, H., & Winner, E. (1984). Words speak louder than actions: Understanding deliberately false remarks. *Child Development, 55,* 1527–1534.

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

Egan, O. (1980). Intonation and meaning. *Journal of Psycholinguistic Research, 9*, 23–39.

Ekman, P. (1985). *Telling lies*. New York: Berkeley Books.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is melody the message? *Child Development, 60,* 1497–1510.

Friend, M. (2000). Developmental changes in sensitivity to vocal paralanguage. *Developmental Science, 3,* 148–162.

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

Frye, D., Zelazo, P. D., & Burack, J. A. (1998). Cognitive complexity and control: I. Theory of mind in typical and atypical development. *Current Directions in Psychological Science, 7,* 116–121.

Furth, S. L., Garg, P. P., Neu, A. M., Hwang, W., Flush, B. A., & Powe, N. R. (2000). Racial differences in access to the kidney transplant waiting list for children and adolescents with end-stage renal disease. *Pediatrics, 106 (4),* 756–761.

Fussell, S., & Moss, L. (eds.) (1998). *Figurative language in emotional communication*. Mahwah, NJ: Lawrence Erlbaum.

Genesee, F., Tucker, G. R., & Lambert, W. E. (1975). Communication skills of bilingual children. *Child Development, 46 (4),* 1010–1014.

Goldie, P. (2002). Emotions, feelings and intentionality. *Phenomenology and the Cognitive Sciences, 1,* 235–254.

Hallstrom, A., Boutin, P., Cobb, L., & Johnson, E. (1993). Socioeconomic status and prediction of ventricular fibrillation survival. *American Journal of Public Health, 83 (2),* 245–248.

Hays, W. (1973). *Statistics for the social sciences*. New York: Holt, Rinehart & Winston.

Hoffmann, C. (1991). *An introduction to bilingualism*. New York: Longman.

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences, 106 (16),* 6556–6560.

Kreuz, R. J. (1996). The use of verbal irony: Cues and constraints. In J. S. Mio & A. N. Katz (eds.), *Metaphor: Implications and applications*, pp. 23–38. Mahwah, NJ: Lawrence Erlbaum.

Kreuz, R. J. (2000). The production and processing of verbal irony. *Metaphor and Symbol, 15,* 99–107.

Kreuz, R. J., & Roberts, R. M. (1995). Two cues for verbal irony: Hyperbole and the ironic tone of voice. *Metaphor and Symbolic Activity, 10,* 21–31.

Labov, W., & Fanshel, D. (1977). *Therapeutic discourse: Psychotherapy as conversation*. New York: Academic Press.

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition, 11 (1),* 81–93.

MathWorks, T. (2007). From http://www.mathworks.com (retrieved August 2007).

Mazzocco, M. M. M. (1999). Developmental changes in indicators that literal interpretations of homonyms are associated with conflict. *Journal of Child Language, 26,* 393–417.

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

Milosky, L. M., & Ford, J. A. (1997). The role of prosody in children's inferences of ironic intent. *Discourse Processes, 23*, 47–61.

Mitchell, W. G., Zhou, Y., Chavez, J. M., & Guzman, B. L. (1992). Reaction time, attention, and impulsivity in epilepsy. *Pediatric Neurology, 8 (1),* 19–24.

Morton, J. B., & Munakata, Y. (2002). Are you listening? Exploring a developmental knowledge–action dissociation in a speech interpretation task. *Developmental Science, 5,* 435–440.

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development, 72,* 834–843.

Morton, J. B., Trehub, S. E., & Zelazo, P. D. (2003). Sources of inflexibility in children's understanding of emotion in speech. *Child Development, 74,* 1857–1868.

Ortony, A. (1975). Why metaphors are necessary and not just nice. *Educational Theory, 25,* 45–53.

Pelli, D. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437.

Rathore, S. S., Masoudi, F. A., Wang, Y., Curtis, J. P., Foody, J. M., Havranek, E. P., & Krumholz, H. M. (2006). Socioeconomic status, treatment, and outcomes among elderly patients hospitalized with heart failure: Findings from the National Heart Failure Project. *American Heart Journal, 152 (2),* 371–378.

Roberts, R. M., & Kreuz, R. J. (1994). Why do people use figurative language? *Psychological Science, 5,* 159–163.

Rockwell, P. (2000). Lower, slower, louder: Vocal cues of sarcasm. *Journal of Psycholinguistic Research, 29,* 483–495.

Saunders, G. (1983). *Bilingual children: Guidance for the family*. Clevedon: Multilingual Matters.

Siegal, M., Iozzi, L., & Surian, L. (2009). Bilingualism and conversational understanding. *Cognition, 110,* 115–122.

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

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Sullivan, K., Winner, E., & Hopfield, N. (1995). How children tell a lie from a joke: The role of second-order mental state attributions. *British Journal of Developmental Psychology, 13,* 191–204.

Vaillancourt, C., Lui, A., De Maio, V. J., Wells, G. A., & Stiell, I. G. (2008). Socioeconomic status influences bystander CPR and survival rates for out-of-hospital cardiac arrest victims. *Resuscitation, 79 (3),* 417–423.

Ward, M. M. (2008). Socioeconomic status and the incidence of ESRD. *American Journal of Kidney Diseases, 51 (4),* 563–572.

Westenberg, P. M., Siebelink, B. M., Warmenhoven, N. J. C., & Treffers, P. D. A. (1999). Separation anxiety and overanxious disorders: Relations to age and level of psychosocial maturity. *Journal of the American Academy of Child & Adolescent Psychiatry, 38 (8),* 1000–1007.

Winner, E., & Leekam, S. (1991). Distinguishing irony from deception: Understanding the speaker's second-order intention. *British Journal of Developmental Psychology, 9,* 257–270.

Yow, W. Q., & Markman, E. M. (2007). Monolingual and bilingual children's use of mutual exclusivity assumption and pragmatic cues in word learning. Poster presented at the meeting of the Society for Research in Child Development, Boston, MA.

Yow, W. Q., & Markman, E. M. (2009). Understanding a speaker's communicative intent: Bilingual children's heightened sensitivity to referential gestures. In J. Chandlee, M. Franchini, S. Lord & G. Rheiner (eds.), *Proceedings of the 33rd Annual Boston University Conference on Language Development* (vol. 2), pp. 646–657.

Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function. *Current Directions in Psychological Science, 7,* 121–126.