CHILD DEVELOPMENT

Child Development, September/October 2011, Volume 82, Number 5, Pages 1648–1660

Children's Judgments of Emotion From Conflicting Cues in Speech: Why 6-Year-Olds Are So Inflexible

Matthew Waxer and J. Bruce Morton
University of Western Ontario

Six-year-old children can judge a speaker's feelings either from content or paralanguage but have difficulty switching the basis of their judgments when these cues conflict. This inflexibility may relate to a lexical bias in 6-year-olds' judgments. Two experiments tested this claim. In Experiment 1, 6-year-olds (n = 40) were as inflexible when switching from paralanguage to content as when switching from content to paralanguage. In Experiment 2, 6-year-olds (n = 32) and adults (n = 32) had more difficulty when switching between conflicting emotion cues than conflicting nonemotional cues. Thus, 6-year-olds' inflexibility appears to be tied to the presence of conflicting emotion cues in speech rather than a bias to judge a speaker's feelings from content.

Young children can accurately judge the emotion of a speaker either from what she says or how she speaks. For example, 4-year-olds accurately judge emotion from descriptions of familiar contexts (Borke, 1971) and accurately label emotional paralanguage when the content is semantically neutral (Baltaxe, 1991; Dimitrovsky, 1964; Morton, Trehub, & Zelazo, 2003), or rendered uninterpretable through filtering (Friend, 2000; Morton & Trehub, 2001). And by 6 years of age, children can use paralinguistic cues to make judgments about basic emotions such as happiness, sadness, and anger (Baltaxe, 1991; Dimitrovsky, 1964; Friend, 2000; Morton & Trehub, 2001).

Despite these basic abilities, 6-year-olds have difficulty flexibly switching the basis of their judgments from propositional to paralinguistic emotion cues when these cues conflict (Friend, 2000; Morton & Munakata, 2002; Morton et al., 2003). In one study, 6-year-old children were presented with utterances that contained conflicting emotion cues (e.g., ''I won a prize,'' spoken sadly), and were instructed to judge a speaker's feelings on the basis of what she said and then switch and judge the speaker's feelings from the sound of her voice (Morton et al., 2003). Although children remembered the instructions, approximately 60% persisted in judging the speaker's feelings from what she said. Very similar findings have been reported in other studies (Eskritt & Lee, 2003; Friend, 2000; Friend & Bryant, 2000; Morton & Trehub, 2001; Solomon & Ali, 1972). The phenomenon is compelling as it falls at the intersection of children's understanding of spoken language, their concepts of mixed emotion, and general age-related constraints on cognitive flexibility.

Current models emphasize the role of language processing and domain general constraints on cognitive flexibility in driving this effect. According to Friend (2000, 2003), for example, children's inflexibility reflects both a language-specific attentional bias and domain-general processing constraints. The attentional bias is thought to play a role in early language development by preferentially weighting the acoustic specification of lexical units in continuous speech. Domain-general constraints by contrast are observed in cases in which propositional and paralinguistic cues to emotion conflict and listeners need to consider both cues when judging a speaker's feelings. Under these conditions, domain-general processing limitations cause children to accord greater attention to one of these two sources of information. By this account, 6-year-olds' inflexibility reflects both domain-specific and domain-general factors.

This research was supported by a National Science and Engineering Research Council Discovery Grant to J.B.M.

Correspondence concerning this article should be addressed to J. Bruce Morton, Department of Psychology and Graduate Programme in Neuroscience, University of Western Ontario, Westminster Hall, London, Ontario, Canada N6A 3K7. Electronic mail may be sent to bmorton3@uwo.ca.

 0009-3920/2011/8205-0023
DOI: 10.1111/j.1467-8624.2011.01624.x

A complementary model views children's inflexibility as a product of the combined influence of distinct active and latent representations (Cepeda & Munakata, 2007; Morton & Munakata, 2002, 2009). As implemented in neural network models, latent representations take the form of connections between units that process a particular feature or dimension of a stimulus whereas active representations take the form of sustained activity in particular processing, or working memory, units. Latent representations change in strength according to a Hebbian learning rule, such that connections between units that show correlated activity become stronger over time (i.e., units that fire together, wire together). Active representations, by contrast, change in strength as a function of recurrent connections that working memory units make with themselves. When these recurrent connections are sufficiently strong, working memory units can form robust active representations of task instructions and efficiently exert top-down control over other parts of the network. Thus, experimentally varying the strength of recurrent connections has been used as a means of formally simulating hypothesized changes in working or ''active'' memory that occur with development. Drawing on these ideas, Morton and Munakata (2002) proposed that the latent representation of propositional content is stronger than that of affective paralanguage, much as the latent representation of word meaning is stronger than that of font color in models of the color–word Stroop interference (Cohen, Dunbar, & McClelland, 1990). This difference reflects the fact that listeners have more experience judging utterances according to message content than by paralanguage, and leads to a ''bias,'' or preference, to respond on the basis of message content among listeners of all ages. Overriding this bias, as is required in different listening contexts or by various task instructions (e.g., ''listen to the speaker's voice and judge how she feels''), requires input from active representations. Older ''adult'' models form strong active representations of contextual cues and task instructions and can therefore easily overcome the bias to respond to content. Younger ''child'' models, by contrast, form weak active representations of task instructions and therefore typically persist in basing their judgments on content despite instructions to do otherwise.

Despite differences in terminology and formal implementation, the two accounts are similar in that they both link children's inflexibility to language-specific biases and domain-general constraints. Consistent with these ideas, young children normally favor propositional over paralinguistic and facial cues when judging a speaker's feelings (Eskritt & Lee, 2003; Friend, 2000; Morton & Trehub, 2001), attitudes (Bugental, Kaswan, & Love, 1970; Solomon & Ali, 1972; Solomon & Yaeger, 1969), and certainty (Furrow, Podrouzek, & Moore, 1990; Moore, Harris, & Patriquin, 1993) from utterances with multiple cues, and are slower to label the identity of a speaker in the presence of semantically conflicting compared to semantically neutral words (Jerger, Martin, & Pirozzolo, 1998). This bias recedes with development (Ackerman, 1986; Morton & Trehub, 2001) as children become more sensitive to the demands of different listening contexts (Morton & Munakata, 2002; Morton et al., 2003) and more effective at inhibiting salient lexical information (Friend, 2000), but can contribute to inflexibility in young children who find it difficult to suppress prepotent responses (Gerstadt, Hong, & Diamond, 1996).

There are, however, several open questions for Friend's (2000) and Morton and Munakata's (2002) accounts. Although children normally exhibit a bias to respond to message content given discrepant emotion cues, the stability of this bias and its role in driving inflexibility has not been carefully examined. There is for example evidence that priming children to attend to paralanguage can quickly reverse their normal response bias (Morton et al., 2003, Experiment 1). In one experiment, 6-year-old children were presented utterances that contained emotionally neutral content (e.g., ''There is water in the pail'') spoken with emotional paralanguage and were asked to judge the speaker's feelings. After several such priming trials, children preferentially responded to the paralanguage rather than the content of utterances with conflicting cues to emotion (Morton et al., 2003, Experiment 1). However, it is not known whether children would be as inflexible when required to switch back to content following such a priming procedure as they are switching from content to paralanguage. Such evidence would suggest that children's inflexibility is tied more to general constraints on cognitive flexibility rather than general constraints *and* language-specific biases (Friend, 2000; Friend & Bryant, 2000).

A second question is whether 6-year-olds' inflexibility stems from a language-specific attentional bias that preferentially weights the acoustic specification of the lexical content of speech (Friend, 2000; Friend & Bryant, 2000) or is tied to normative beliefs about spoken propositions. Indeed, in most experiments, children are asked to infer a speaker's feelings from spoken descriptions of familiar emotive situations (e.g., ''I got ice cream for dessert today''). Although children obviously attend to the constituent words and phrases of the utterances, their inferences about the speaker's feelings may be based on normative beliefs about speakers (e.g., speakers normally say things that are true; people normally like ice cream), objects (e.g., ice cream normally tastes good), and events (e.g., receiving something good normally makes a person happy; Terwogt & Rieffe, 2003). Switching away from content therefore might be difficult for children because doing so would require abandoning strongly held normative beliefs about speakers and the world. On this account, children should exhibit greater flexibility were emotion cues presented lexically rather than propositionally. Although children could formulate a response of happy or sad to words that carry emotional meaning (e.g., *loss*), doing so would not invite the use of normative beliefs about speakers and states of the world in the same way that spoken propositions do. If, on the other hand, children's inflexibility is due in part to selective attention to word meaning, then children should show equivalent inflexibility for both propositions and emotion words in isolation.

We tested these alternative accounts of children's inflexibility in Experiment 1 through targeted modifications of the Speech Interpretation Task (SIT; Morton et al., 2003), a task in which children are presented a series of utterances with discrepant emotion cues (e.g., ''I won a prize'' spoken with sad paralanguage) and are asked first to respond on the basis of message content and then switch and respond on the basis of paralanguage. To investigate whether children's inflexibility relates to difficulty overcoming a bias to content specifically or difficulty overcoming bias more generally, we compared performance in a version of the SIT that required a switch from content to paralanguage with a version that required a switch from paralanguage to content. Accounts that attribute inflexibility to a language-specific bias predict that the paralanguage-to-content switch would be easier than the content-to-paralanguage switch, as the former involves switching to a preferred task whereas the later involves switching from a preferred task (Friend, 2000; Friend & Bryant, 2000; Morton & Munakata, 2002). By contrast, accounts that tie children's inflexibility to more general developmental constraints on flexible thinking would predict that both switches would be difficult (Deák, 2000; Friend, 2000; Friend & Bryant, 2000; Zelazo & Frye, 1997). To investigate whether children's inflexibility is related to the use of propositional utterances, we presented conflicting emotion cues both lexically and propositionally. If children are inflexible in their judgments of spoken utterances because they rely on normative beliefs about speakers and their propositions, then in principle it should be easier for children to switch the basis of their judgments given lexical rather than propositional content. Child listeners would presumably be less likely to rely on normative beliefs when judging the emotional meaning of a single word than a spoken proposition (Terwogt & Rieffe, 2003).

Experiment 1

Method

Participants. Forty (20 boys) 6-year-old children (M = 6.2, range = 6.0–6.5) participated. Participants were predominantly from Caucasian middle-class families and were recruited through the use of birth announcements. All participants spoke English fluently.

Apparatus and stimuli. Children were tested individually by a computer in a quiet room. Children sat facing the computer monitor and the experimenter sat beside them. The experimenter used the computer keyboard to call for trials, and children responded using a button box connected to the computer. One button, labeled the ''happy button,'' had a drawing of a happy face on it, and the other button, labeled the ''sad button,'' had a drawing of a sad face on it.

Happy, sad, and neutral lexical and propositional content was variously combined with happy, sad, and neutral paralanguage to create a set of 56 stimuli, including 28 propositional and 28 single-word utterances. Ten propositions and 10 words had happy content (6 each spoken with sad paralanguage and 4 each spoken with neutral paralanguage), 10 propositions and 10 words had sad content (6 each spoken with happy content and 4 each spoken with neutral paralanguage), and 8 propositions and 8 words had neutral content (4 each spoken with happy paralanguage and 4 each spoken with sad paralanguage). All utterances and words were spoken by the same woman and were digitally recorded on a computer using GoldWave and Adobe Audition software. The propositional content used in the utterances (see Table 1) was adopted from previous studies (Morton & Trehub, 2001; Morton et al., 2003) and was readily interpretable to 6-year-olds. Words for the lexical task (see Table 2) were generated using the MRC Psycholinguistic Database (Coltheart, 1981) and were selected if they were rated as comprehensible by

Table 1
Neutral and Emotional Propositions

Sad propositional content	
1.	I am not allowed to go swimming with my friends.
2.	I dropped my ice cream cone.
3.	I lost my stamp collection.
4.	All the kids at school tease me.
5.	I lost my baseball glove today.
6.	My best friend doesn't like me anymore.
7.	I lost the toy my grandma gave me for Christmas.
8.	I lost all my money on the way to the store.
9.	I fell of my bike, and everyone made fun of me.
10.	I am not allowed to go outside and play with my friends.
Happy propositional content	
1.	My baseball team won the game.
2.	Grandpa gave me some candy for being good.
3.	My coach says that I am the fastest runner on the team.
4.	Grandma gave me a video game for Christmas.
5.	My mommy gave me a treat.
6.	I had my favorite cake for dessert.
7.	Grandma told me that I am very special.
8.	My soccer team just won the championship.
9.	Dad gave me a new bike for my birthday.
10.	I got an ice cream for being good.
Emotionally neutral propositional content	
1.	My dad is wearing his glasses.
2.	I made a sweater out of wool.
3.	I put my marbles in a bag.
4.	My daddy drove his car down the street.
5.	I sat down on the chair.
6.	I carried water in a pail.
7.	I am using the hose.
8.	My doll is wearing a dress.

Table 2
Neutral and Emotional Words

Sad words	
1.	Afraid
2.	Alone
3.	Bad
4.	Cry
5.	Lonely
6.	Lost
7.	Mean
8.	Sad
9.	Worry
10.	Wrong
Happy words	
1.	Awesome
2.	Brave
3.	Friend
4.	Good
5.	Happy
6.	Laugh
7.	Perfect
8.	Proud
9.	Share
10.	Smile
Emotionally neutral words	
1.	Chair
2.	Pen
3.	Stone
4.	Table
5.	Tape
6.	Tree
7.	Water
8.	Window

4-year-olds. Utterances and words were presented at a comfortable listening level, in an order that was randomized for each participant.

Tasks and procedures. There were four tasks altogether, including two content-to-paralanguage switch tasks (one lexical and one propositional) and two paralanguage-to-content switch tasks (one lexical and one propositional). All tasks included eight priming trials, six preswitch trials, and six postswitch trials. Priming trials preceded preswitch trials and were used to orient participants to the appropriate dimension of speech. In the content-to-paralanguage tasks, priming trials consisted of emotional content (lexical in the lexical task, propositional in the propositional task) spoken with neutral paralanguage, and children were instructed to listen to what the speaker said and judge whether she was feeling happy or sad. In the paralanguage-to-content tasks, priming trials consisted of neutral content (lexical in the lexical task, propositional in the propositional task) spoken with emotive paralanguage and children were instructed to listen to the sound of the speaker's voice and judge whether she was feeling happy or sad. Preswitch trials followed priming trials immediately without delay and differed from priming trials only insofar as the stimuli contained conflicting emotion cues. After completing the priming and preswitch trials, children were stopped and asked to play a new game. In the content-to-paralanguage switch, they were told that in the new game they were not to listen to what the speaker was saying but to how her voice sounded. To clarify the meaning of listening to someone's voice, and to ensure that children were capable of labeling examples of emotional paralanguage, the experimenter engaged each participant in a brief dialogue about happy and sad voices and had the children label a few hummed examples. In the paralanguage-to-content switch, they were told that they were not to listen to how the speaker's voice sounded but to what she was saying. To clarify the notion of listening to what someone was saying, the experimenter engaged each child in a short dialogue about happy and sad words and had children provide examples. The instructions were then repeated and the postswitch trials administered.

Tasks were administered in one of four experimental conditions. In all conditions, participants were administered two tasks, one with lexical content and one with propositional content. In Conditions 1 and 2, both tasks were paralanguage-to-content switch tasks. In Conditions 3 and 4, both tasks were content-to-paralanguage switch tasks. Conditions 1 and 2 and Conditions 3 and 4 differed only in terms of the order of task administration (i.e., lexical vs. propositional task first). Children were randomly assigned to conditions, resulting in four groups of identical gender ratio and that did not significantly differ in age (see Table 3). Instructions in all four conditions were identical, and only differed as a function of task order and switch type.

Results

Priming and preswitch trials. There was no difference in priming or preswitch performance across conditions and tasks. Participants were highly accurate, reaching or surpassing a criterion of at least 80% correct across both types of trials in all conditions and tasks. All children were therefore included in the final analysis.

Postswitch trials. Postswitch scores were based on the number of correct postswitch trials and ranged from 0 (*correct on zero trials*) to 6 (*correct on six trials*). Most children either switched on every postswitch trial or failed to switch on any postswitch trial. Scores were therefore nonnormally distributed with approximately 55% of scores taking the value of 0 and a further 25% taking the value of 6 in all conditions.

Although the primary interest in the analysis was whether postswitch performance differed across stimulus types (i.e., lexical vs. propositional) and switch types (i.e., content to paralanguage vs. paralanguage to content), the first step was to test for possible effects of task order. Therefore, a 2 (stimulus type) × 2 (switch type) × 2 (task order) mixed analysis of variance (ANOVA) was conducted. There was no effect of task order and no interactions between task order and any of the other two factors. Therefore, the data were collapsed across order. As shown in Table 3, mean postswitch scores across the four tasks were quite comparable and ranged between 1.8 and 2.6. At face value, these scores suggest that children performed at chance. However, this was not the case, as individual children performed systematically, as mentioned above. A 2 (stimulus type) × 2 (switch type) mixed ANOVA confirmed that there was no effect of stimulus type or switch type, and no interaction of these factors.

Table 3
Mean (Standard Deviation) Number of Correct Postswitch Trials as a Function of Switch Version and Stimulus Type, Experiment 1

	Stimulus type	
Version	Propositional	Lexical
P → C	1.8 (2.7)	2.1 (2.6)
C → P	2.5 (2.8)	2.6 (2.7)

Note. P → C = paralanguage to content; C → P = content to paralanguage.

Discussion

Consistent with previous studies (Friend, 2000; Morton et al., 2003), a majority of 6-year-old children persisted in basing their judgments of a speaker's feelings on propositional content when instructed to switch to paralanguage. These results were not attributable to difficulty correctly labeling the emotional paralanguage or selectively responding to the paralanguage in the presence of conflicting content as all children in the paralanguage to content switch conditions performed at ceiling in priming and preswitch trials. However, children were equally inflexible when instructed to switch their attention from paralanguage to content, and this was true regardless of whether content was conveyed propositionally or lexically. The latter finding suggests that children's persistent use of content in the face of instructions to switch to paralanguage is not tied to difficulty overriding normative beliefs about speakers and the events they describe. Children were presumably more likely to appeal to these beliefs when formulating a response to spoken propositions than single words but were no less likely to switch under these circumstances.

The fact that the lexical bias was easily modified by simple priming procedures and that children were as inflexible when switching to content as they were when switching to paralanguage suggests that 6-year-old children's inflexibility is not related to difficulty overcoming a bias to content (Friend, 2000; Morton & Munakata, 2002) but to more general age-related constraints on cognitive flexibility (Friend, 2000; Friend & Bryant, 2000). Perseverative inflexibility, like that observed in Experiment 1, has been linked for example to constraints in the representation and use of higher order rules. On this account, children can represent pairs of preswitch and postswitch rules but have difficulty switching between them because they are unable to represent a higher order rule. Higher order rules support the recognition that preswitch and postswitch rules specify opposite responses to the same stimuli. Alternatively, children may have difficulty redescribing utterances in a new way having previously described them a different way. By this account, perseverative inflexibility like that observed in Experiment 1 reflects a problem acknowledging that multiple contradictory descriptions can be simultaneously applied to the same stimulus (Kloo, Perner, Kerschuber, Dabernig, & Aichorn, 2008).

It is possible though that 6-year-old children's inflexibility is specifically related to the presence of conflicting emotional stimuli in the SIT. After all, 6-year-old children have little difficulty flexibly sorting objects by color and shape (e.g., Zelazo, Frye, & Rapus, 1996), suggesting that children this age possess the basic executive skills necessary for considering multiple interpretations of a single stimulus. Implementing these skills in the context of problems involving conflicting emotional stimuli may however be uniquely difficult, perhaps because of the immature status of children's concepts of mixed emotions. Although it is within the capacity of young infants to experience mixed emotions—ambivalent infants, for example, express both relief and anger toward the mother during the reunion episode of the Strange Situation procedure (Bretherton & Waters, 1985; although see Larson, To, & Fireman, 2007)—it is not until much later in development that children fully recognize and understand mixed emotion states either in themselves or others (Harris, 1983; Harter & Buddin, 1987; Lane & Schwartz, 1987; Larson et al., 2007). In one study, children were presented with short vignettes in which a character's lost dog returns home with a cut ear. When asked how the character felt, 6-year-old children either responded happy, appealing to the fact the dog had come home, or sad, appealing to the dog's injury. Only 10-year-old children maintained that the character would feel happy and sad (Harris, 1983, Experiment 2). Advances in children's understanding of mixed emotions in others parallels changes in their understanding of their own mixed emotions. As early as 7 years of age, children recognize that they can simultaneously experience two positive or two negative feelings in response to an event (e.g., ''I felt sad and mad when my brother hit me''), but only by late childhood or early adolescence do they admit to being able to feel both positively and negatively about one situation (e.g., ''I felt happy to get a bike for my birthday, but sad it wasn't a 10-speed''). Insight into mixed emotion states in oneself and others is thought to be predicated on conceptual changes that occur in childhood (Arsenio & Lover, 1995; Donaldson & Westerman, 1986; Harter & Buddin, 1987; Lane & Schwartz, 1987; Larson et al., 2007) and may have implications for 6-year-olds' ability to recognize simultaneous emotional meanings in spoken words or phrases.

Switching between conflicting emotional interpretations of the same utterance may however be difficult for children because of the nature of basic emotion concepts like happiness and sadness. Basic emotion concepts have been likened to nodes in a distributed semantic network around which are collected descriptions of prototypical evocative situations, verbal labels, and specific autonomic reactions (Bower, 1981; Lang, 1994). Nodes become activated by both symbolic and physiological means and, when activated above some threshold, transmit excitation to other nodes representing features or behaviors commonly assigned to that emotion, and reciprocally inhibit emotions of opposing quality. On this account, concepts of happiness and sadness represent distinct and reciprocally inhibiting nodes within a distributed semantic space, such that increasing activation of one concept leads to a suppression of activity in the other. Flexibly applying these concepts to the same utterance would therefore require the resolution of conflict between the opposing members of the pair and be highly demanding of top-down executive control.

Experiment 2

The goal of Experiment 2 was to disentangle these alternative explanations. Six-year-old children and adults were administered three modified versions of the SIT that all required a shift in responses to different dimensions of the same spoken stimuli, but differed in terms of the stimuli used. One version, termed the Emotion-SIT, was identical to the lexical task used in Experiment 1 and required that participants switch between dimensions of spoken words that were conflicting and emotional. A second version, termed the Opposites-SIT, used nonemotional words spoken with opposite but nonemotional paralanguage (e.g., the word *high* spoken in a low pitch). This task required that participants switch between dimensions of spoken words that were conflicting but nonemotional. In a third task, termed the Basic-SIT, participants were presented words spoken by two different speakers, and had to switch from responding on the basis of speaker identity to word meaning or vice versa. In this case, participants had to switch between dimensions of spoken words that were nonconflicting and nonemotional.

Our predictions were as follows. If 6-year-olds inflexibility relates to general constraints on the representation and use of higher order rules, then children should perform comparably in all three tasks, as all tasks involve switching between pairs of contradistinctive preswitch and postswitch rules. If, however, children's inflexibility relates to difficulty simultaneously applying contradictory descriptions to the same stimuli (Kloo et al., 2008), then they should perform better in the Basic- than in the Emotion- and Opposites-SIT, as only the Emotion- and Opposites-SIT involved conceptually contradictory stimuli. If children's inflexibility is tied specifically to the presence of emotional stimuli, then they should perform better in the Basic- and Opposites-SIT than in the Emotion-SIT. Finally, an examination of age-related differences in the pattern of performance across tasks would clarify the role of concepts of mixed emotions in the Emotion-SIT. For example, if children's inflexibility were confined to the Emotion-SIT, but adults, who presumably have mastered concepts of mixed emotions, found all three tasks relatively easy, this would suggest that children's inflexibility is tied to constraints in their understanding of mixed emotions. If, however, children and adults had more difficulty with the Emotion-SIT than the Basic- and Opposites-SIT, this would imply that resolving emotional conflict is difficult regardless of whether one possesses concepts of mixed emotion.

Method

Participants. Participants included 32 children (16 males) and 32 young adults (15 males). The children were all 6 years old ($M = 6.2$ years, range = 6.0–6.4), and the adults ranged in age from 17 to 21 years ($M = 17.4$ years). Children were predominantly from Caucasian middle-class families and were recruited through the use of birth announcements; adults were students enrolled in introductory psychology courses who participated in exchange for course credit. Adults provided written consent to their participation. Parents provided written consent for their children's participation. All participants spoke English fluently.

Apparatus and stimuli. Testing took place in the same room and with the same equipment as that used in Experiment 1. Participants responded using a button box connected to a computer. Pictures affixed to the buttons differed for each task. For the Emotion-SIT, pictures consisted of a happy face and a sad face. For the Opposites-SIT, pictures consisted of an arrow pointing up and an arrow pointing down. For the Basic-SIT, pictures consisted of a boy with a book on his shirt and picture of a girl with a tree on her shirt.

Stimuli for the Emotion-SIT were the same as that used in Experiment 2 (see Table 2). Stimuli for the Opposites-SIT consisted of 14 words (see Table 4). Ten words were both nondirectional and nonemotional in meaning (e.g., *car*), and included four spoken in a neutral pitch, three spoken in a high pitch, and three spoken in a low pitch. Two words had meanings associated with ''up'' and were spoken in a low pitch, and two words had meanings associated with ''down'' and were spoken in a high pitch. Stimuli for the Basic-SIT (see Table 4) consisted of two words, the word *tree* spoken by a male and the word *book* spoken by a female. Thus, for each word, the meaning matched one target and the gender of the speaker matched the other.

All stimuli were digitally recorded via a personal computer using GoldWave and Adobe Audition software. Words for all three tasks were spoken by the same woman. Stimuli used in Opposites-SIT were initially recorded in a neutral pitch. High- and low-pitch versions were then created using Adobe Audition software with a procedure that

Table 4
Experimental Stimuli for the Opposites- and Basic-SIT Tasks

Opposites-SIT neutral words
1. Ball
2. Barn
3. Car
4. Can
5. Hill
6. Truck
Opposites-SIT directional words
1. Down
2. High
3. Low
4. Up
Basic-SIT words
1. Book
2. Tree

preserved the speaking rate. Words were selected from the MRC Psycholinguistic Database (1987) and were all comprehensible to 4-year-old children. All stimuli were presented at a comfortable listening level. The order of stimuli presentation was randomized for each participant.

Procedure. There were two versions of all three tasks, content to voice sound and voice sound to content, that differed only in the order of operations. Procedures for the two versions of the Emotion-SIT were identical to those in Experiment 1.

Both versions of the Opposites-SIT consisted of priming, preswitch, and postswitch blocks with each block consisting of six trials. Prior to the priming trials of the content to voice-sound version, participants instructed to respond on the basis of what the speaker said. When she said a high word, they were to press the ''up button'' (i.e., the button marked with an arrow pointing up) and when she said a low word, they were to press the ''down button'' (i.e., the button with an arrow pointing down). To ensure children understood the instructions, the experimenter briefly explained the difference between ''high'' and ''low'' words and had children offer examples. All children generated appropriate examples. Participants then completed six priming trials in which directional words were spoken in a neutral pitch, and six preswitch trials in which directional words were spoken with the opposite voice (e.g., the word ''low'' spoken with high pitch). Preswitch trials followed the priming trials without interruption. After completing the preswitch trials, participants were stopped and provided with instructions to switch and now listen to how the speaker's voice sounded, not to what she was saying. If the speaker spoke in a high pitch, they were to press the ''up button'' and if she spoke in a low pitch they were to press the ''down button.'' To ensure children understood the instructions, the experimenter briefly explained the difference between high- and low-pitch voices and had children label hummed examples of high- and low-pitch voices. All children correctly labeled the experimenter's examples. Instructions for the voice-sound to content version were identical except that the instructions about listening to content came prior to postswitch trials and instructions about listening to voice sound came prior to the priming trials.

Both versions of the Basic-SIT consisted of six preswitch and six postswitch trials, but no priming trials due to difficulty generating a gender-neutral voice. In the content to voice-sound version, participants were told that they were going to be listening to two different speakers. Pointing to the corresponding pictures affixed to the button box, the experimenter introduced Sarah, who had a picture of a tree on her shirt, and Mike, who had a picture of a book on his shirt. Participants were told they had to listen to what the speakers said in order to get the answer right. When they hear the word *book*, they had to press the picture of the person with the book on their shirt and when they heard the word *tree* they had to press the picture of the person with the tree on their shirt. To ensure that the children understood the instructions, the experimenter asked each child to reiterate the instructions and indicate which button they would press for each word, which all children did correctly.

Following the preswitch trials, participants were told to switch and listen to who was speaking, not to what they were saying. Participants were instructed to press the picture of Mike when they heard Mike speaking and to press the picture of Sarah when they heard Sarah speaking. To ensure children understood the instructions, the experimenter explained the difference between male and female voices and reiterated that they were to listen to who was speaking and not to what was being said.

Instructions for the voice-sound to content version were identical except for their order or presentation.

All participants were tested individually and were randomly assigned to one of two experimental conditions. In both conditions, participants completed all three tasks. However, in one condition, participants switched from content to the sound of the utterance for all tasks, while in a second condition, they switched from the sound of the utterance to the content for all tasks.

Results

Priming and preswitch trials. Participants had to achieve a criterion of five of six trials correct in both the priming and the preswitch trials in order to be included in the analysis, except in the Basic-SIT for which there were no priming trials. All participants performed at ceiling or made one mistake in priming and preswitch trials of both content-to-voice and voice-to-content versions of all three tasks. Therefore, all participants were included in the analysis.

Postswitch trials: Accuracy. Postswitch scores (shown in Table 5) were based on the number of correct postswitch responses and ranged in value from 0 to 6, with approximately 45% and 35% of children achieving a score of 0 and 6, respectively, and 0% and 85% of adults achieving a score of 0 and 6, respectively, in all three tasks. Postswitch performance in the Opposite- and Basic-SIT was markedly better than in the Emotion-SIT for the children, while postswitch performance in the three tasks did not differ for the adults. Additionally, there were no consistent differences across versions (paralanguage to content vs. content to paralanguage). A 3 (task) × 2 (version) × 2 (age group) repeated measures ANOVA confirmed effects of task, $F(2, 120) = 41.3$, $p < .001$, $\eta^2 = .41$; age group, $F(1, 60) = 34.9$, $p < .001$, $\eta^2 = .37$; and an interaction between task and age group, $F(2, 120) = 28.5$, $p < .001$, $\eta^2 = .32$. To identify the source of the interaction, repeated measures ANOVA's on the effect of task were run separately for children and adults. For children, there was an effect of task, $F(2, 62) = 37.16$, $p < .001$, with Bonferroni-corrected post hoc contrasts indicating fewer correct postswitch responses in the Emotion- than in the Opposites-SIT, $t(31) = -5.85$, $p < .001$, and the Basic-SIT, $t(31) = -7.71$, $p < .001$, and no difference in the number of correct postswitch responses in the Opposite- and Basic-SIT, $t(31) = -2.11$, *ns*. For adults, there was no effect of task, $F(2, 62) = 2.03$, *ns*.

Switch costs. Because adults performed the three tasks with comparable accuracy, we compared adults' switch costs across the three tasks. For each task, switch cost was computed as the difference in mean response time of correct preswitch and postswitch trials. Mean switch costs (milliseconds) and standard deviations for content-to-voice and voice-to-content versions of the Basic-, Opposites-, and Emotion-SIT tasks are shown in Table 6.

Switch costs were markedly greater in the Emotion-SIT than in either the Opposites- or Basic-SIT, and there were no consistent differences across versions (paralanguage to content vs. content to paralanguage) of the three tasks. A 3 (task) × 2 (version) repeated measured ANOVA confirmed an effect of task, $F(2, 60) = 10.4$, $p < .001$, $\eta^2 = .39$, and there was no effect of version, $F(1, 30) = 0.01$, *ns* and no Task × Version interaction, $F(2, 60) = 0.04$, *ns*. Bonferroni-corrected post hoc contrasts of the effect of task indicated that switch costs were greater for the Emotion- than for both the Opposites-SIT, $t(31) = 4.25$, $p < .001$, and the Basic-SIT, $t(31) = 2.93$, $p < .01$.

Table 5
Mean (Standard Deviation) Postswitch Scores Out of 6 as a Function of Task and Version, Experiment 2

		Task		
	Version	Basic-SIT	Opposites-SIT	Emotion-SIT
6-year-olds	P → C	5.3 (1.5)	4.3 (2.6)	2.3 (2.8)
	C → P	5.6 (1.5)	5.0 (2.1)	1.1 (2.4)
Adults	P → C	5.81 (0.544)	5.81 (0.403)	5.50 (0.730)
	C → P	5.75 (0.577)	5.81 (0.544)	5.44 (0.892)

Note. SIT = Speech Interpretation Task; P → C = paralanguage to content; C → P = content to paralanguage.

Discussion

As in Experiment 1, 6-year-old children had difficulty switching between different dimensions of spoken words when those dimensions conveyed conflicting emotional meanings, and this was true whether children were switching from paralanguage to content or from content to paralanguage. Interestingly, the very same children found it comparatively easy to switch between different dimensions of nonemotional speech, even if the dimensions were opposite in meaning. As in Experiment 1, difficulty switching to paralanguage in the Emotion-SIT could not be attributed to difficulty labeling paralanguage or difficulty selectively attending to paralanguage in the presence of conflicting content as children performed at ceiling in the priming and preswitch trials of the paralanguage-to-content switch task. Like children, adults had more difficulty with the Emotion-SIT than with either the Basic- or the Opposites-SIT, as reflected in larger switch costs (in terms of response time) for the former than the latter two tasks. Taken together, the findings help to clarify why 6-year-old children are often inflexible in their judgments of conflicting emotions in speech.

One possibility we considered is that children's inflexibility reflects general age-related constraints on executive control, such as constraints on the representation and use of higher order rules (Zelazo & Frye, 1997), or the capacity for stimulus

Table 6
Mean (Standard Deviation) Switch Costs in Milliseconds as a Function of Task and Version, Experiment 2

		Task		
	Version	Basic-SIT	Opposites-SIT	Emotion-SIT
Adults	P → C	50.31 (22.02)	45.65 (22.64)	81.26 (51.98)
	C → P	49.95 (23.21)	46.08 (26.31)	77.80 (61.89)

Note. SIT = Speech Interpretation Task; P → C = paralanguage to content; C → P = content to paralanguage.

redescription (Kloo et al., 2008). Although the current results do not rule out these explanations, they do highlight certain insufficiencies. For example, participants in all tasks had to coordinate contradistinctive pairs of preswitch and postswitch rules, and in the Opposites- and Emotion-SIT, switch between contradictory descriptions of the same stimuli. However, in contrast to the predictions of these accounts, both the Basic- and Opposites-SIT were easier for children than the Emotion-SIT. This does not imply that higher order rule use or stimulus redescription are not critical aspects of executive control but simply that they are not sufficient explanations for children's inflexibility in the Emotion-SIT. Something related to the presence of conflicting emotion cues in the stimuli made the Emotion-SIT considerably more difficult for children than the Basic- and Opposites-SIT. We will return to the issue of how the present findings might be accommodated by other models of executive control in the General Discussion.

A second possibility we considered was that 6-year-olds' inflexibility in the Emotion-SIT reflects constraints in their understanding of mixed emotions. Insight into mixed emotion states develops gradually, and only by late childhood or early adolescence do children admit that it is possible to simultaneously experience both positive and negative feelings about a single event. These developments are thought to be systematic and stage-like, and predicated on cognitive changes that occur in middle and late childhood. It seemed reasonable therefore to hypothesize that in light of these constraints, it might be difficult for 6-year-old children to acknowledge for example that a speaker feels happy (as indicated by her paralanguage) about a situation that is typically associated with sadness (e.g., loss). In contrast to this prediction, adults who had presumably mastered concepts of mixed emotion states nevertheless showed larger switch costs in the Emotion-SIT than in the Basic- and Opposites-SIT. These data suggest that a mastery of concepts of mixed emotions is not sufficient to abolish the unique difficulties associated with switching attention between conflicting cues to emotion in speech. Of course, this does not rule out the possibility that 6-year-olds' difficulties with the Emotion-SIT stem in part from conceptual rather than processing constraints. Indeed, it would not be surprising that concepts that are difficult to master initially impose a processing burden later. Clarifying these issues is an important avenue for future research.

In our view though, the findings are most consistent with the idea that opposite emotion concepts, such as happiness and sadness, are strongly and mutually inhibitory. Basic emotion concepts have been likened to nodes in a distributed semantic network (Bower, 1981; Lang, 1994). When activated above threshold, nodes are thought to transmit excitation to semantically related nodes but reciprocally inhibit nodes representing opposite concepts. By this account, happiness and sadness reciprocally inhibit one another, such that activity in one node suppresses activity in the other. Flexibly applying these concepts to the same utterance would therefore require the resolution of conflict between the opposing members of the pair and be highly demanding both for children and adults. Switching between opposite concepts like high and low may be easier because inhibitory connections between these concepts may be weaker due to their being less familiar, certainly as applied to sound as was the case in this experiment. Clarifying these issues will require additional research.

General Discussion

Although 6-year-old children are proficient in judging basic emotions from propositional and paralinguistic cues in speech, they are remarkably inflexible in their judgments. Children at this age typically judge a speaker's feelings from message content and persist in doing so even when given explicit instructions to use paralanguage as a basis for their judgments. One account of this phenomenon proposes that early in development, listeners attend selectively to those aspects of speech that are most relevant to comprehension, leading to a lexical bias in the processing of spoken language. This language-specific bias is thought to influence the processing of all forms of spoken language, emotional, and nonemotional alike, and recede in its overt expression with advances in domain-general processes that support higher level attentional control (Friend, 2000). A related account characterizes performance in tasks like the SIT as a competition between active and latent representations (Morton & Munakata, 2002). Implemented in the form of a neural network model, listeners' lexical bias is reflected in stronger connections between units that process content and weaker connections between units that process paralanguage. This difference in connection weights, in turn, causes the model to preferentially respond to content when presented with conflicting propositional and paralinguistic cues to emotion. Overcoming this bias is made possible by the influence of active memory units that maintain a representation of task instructions and modulate competition between underlying latent representations. Despite their differences, both models suggest that children's inflexibility is linked, at least in part, to a prepotent bias to judge a speaker's feelings on the basis of content.

Unexpectedly, children in Experiment 1 were as inflexible when switching from paralanguage to content as they were when switching from content to paralanguage and this was true whether stimuli were lexical or propositional. These results suggest that whatever the underlying basis of children's lexical bias might be, the bias itself is not a source of 6-year-old children's inflexibility. And in Experiment 2, children's inflexibility was confined to a task involving conflicting emotion cues. Other switching tasks involving nonemotional stimuli were comparatively easy for children, even if they involved contradictory stimuli. Importantly, the unique difficulty of switching between alternative contradictory emotional interpretations of spoken words was also evident in adults' performance. Adults showed larger costs when switching responses to emotionally contradictory utterances than to nonemotional utterances, even if those utterances afforded contradictory interpretations. Thus, mastery of concepts of mixed emotion by adults was not sufficient to abolish inflexibility that is so readily observed in 6-year-old children.

Taken together, the present results suggest that the efficacy of domain-general executive processes like switching can vary as a function of the content domain to which they are applied. Concepts of happiness and sadness, for example, may represent distinct and reciprocally inhibiting nodes within a distributed semantic space, such that increasing activation of one concept leads to a suppression of activity in the other. Flexibly applying these concepts to the same utterance would therefore require the resolution of conflict between opposing members of the pair. Flexibly applying concepts like high and low may be easier if one assumes that these concepts are not as strongly associated with paralinguistic pitch variations as is the case with happy and sad. On this assumption, switching between concepts of high and low in the SIT involves less conflict than switching between concepts of happy and sad.

The idea that greater conflict is associated with greater demands on executive control is central to many accounts of executive functioning. According to Diamond (2002), for example, executive functioning tasks demand working memory and inhibitory control and become more demanding in the face of increasing conflict. In the Day-Night Stroop task, for example, children are presented pairs of images and label one ''day'' and the other ''night.'' The task is relatively easy when the pictures are of unfamiliar geometric shapes, but is difficult when children are required to say ''day'' to a picture of the moon and ''night'' to a picture of the sun. The difference relates to the fact that in the latter version, there is conflict between the responses specified by the rule and the prepotent responses to say ''day'' and ''night'' to images of the sun and moon, respectively. The increase in conflict leads to a greater demand on inhibitory control and poorer performance. Similarly, Morton and Munakata (2002) suggest that the reason formally similar tasks such as the DCCS and the SIT are mastered at different ages relates to differences in the amount of conflict involved in these respective tasks. They propose that the SIT is more difficult than the DCCS because children come into the SIT with a strong prepotent bias to respond to the content of spoken utterances. Although the importance of the lexical bias for children's inflexibility was not borne out by the present studies, the general idea that increasing conflict is associated with increasing demands on executive control tasks is consistent with the present findings. How this sort of conflict might be formally implemented deserves future consideration.

While the current findings argue against a role for language-specific biases in 6-year-olds' inflexible judgments of emotion in speech, they do not diminish the potential relevance of these biases for understanding the processing of spoken language and its development. Although a lexical bias in emotional judgments of spoken utterances can be easily displaced by priming procedures and task instructions (at least in adults), the bias itself is readily observed in participants of all ages if probed more subtly, for instance by means of interference effects in auditory Stroop paradigms (Jerger et al., 1998). The origins of the bias and its importance for language development remain important questions for future research. It is possible that the bias simply reflects the consequence of repeated experiences listening to and interpreting spoken utterances. At the same time, neuroconstructivist accounts of development argue that emerging cognitive specialization reflects an interaction of early biases in information processing and experience (Johnson, 2001). It is possible, then, that children's and adults' lexical bias in the processing of spoken language represents a vestige of an earlier bias that plays an important role in the emergence of language.

References

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

Arsenio, W. F., & Lover, A. (1995). Children's conceptions of socio-moral affect: Happy victimizers, mixed emotions and other expectancies. In M. Killen & D. Hart (Eds.), *Morality in everyday life: Developmental perspectives* (pp. 87–128). Cambridge, UK: Cambridge University Press.

Baltaxe, C. A. M. (1991). Vocal communication of affect and its perception in three- to four-year-old children. *Perceptual and Motor Skills*, *72*, 1187–1202.

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

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.

Bretherton, I., & Waters, E. (1985). Growing points of attachment theory and research. *Monographs of the Society for Research in Child Development*, *50* (Serial No. 209), 3–35.

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

Cepeda, N. J., & Munakata, Y. (2007). Why do children perseverate when they seem to know better: Graded working memory or directed inhibition? *Psychonomic Bulletin & Review*, *14*, 1058–1065.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*, 332–361.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33A*, 497–505.

Deák, G. O. (2000). The growth of flexible problem-solving: Preschool children use changing verbal cues to infer multiple word meanings. *Journal of Cognition and Development*, *1*, 157–192.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). Oxford, UK: Oxford University Press.

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

Donaldson, S. K., & Westerman, M. A. (1986). Development of children's understanding of ambivalence and causal theories of emotion. *Developmental Psychology*, *22*, 655–662.

Eskritt, M., & Lee, K. (2003). Do actions speak louder than words? Preschool children's use of the verbal-nonverbal consistency principle during inconsistent communications. *Journal of Nonverbal Behavior*, *27*, 25–41.

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

Friend, M. (2003). What should I do? Language, paralanguage, and behavior regulation in early childhood. *Journal of Cognition and Development*, *4*, 161–183.

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

Furrow, D., Podrouzek, W., & Moore, C. (1990). The acoustical analysis of children's use of prosody in assertive and directive contexts. *First Language*, *10*, 37–49.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1996). The relationship between cognition and action: Performance on children 3 ½-7 years old on a Stroop-like day-night test. *Cognition*, *53*, 129–153.

Harris, P. L. (1983). Children's understanding of the link between situation and emotion. *Journal of Experimental Child Psychology*, *36*, 490–509.

Harter, S., & Buddin, B. J. (1987). Children's understanding of the simultaneity of two emotions: A five-stage developmental acquisition sequence. *Developmental Psychology*, *23*, 388–399.

Jerger, S., Martin, R. C., & Pirozzolo, F. J. (1998). A developmental study of the auditory Stroop effect. *Brain and Language*, *35*, 86–104.

Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews, Neuroscience*, *2*, 475–483.

Kloo, D., Perner, J., Kerschuber, A., Dabernig, S., & Aichorn, M. (2008). Sorting between dimensions: Conditions of cognitive flexibility in preschoolers. *Journal of Experimental Child Psychology*, *100*, 115–134.

Lane, R. D., & Schwartz, G. E. (1987). Levels of emotional awareness: A cognitive-developmental theory and its application to psychopathology. *American Journal of Psychiatry*, *144*, 133–143.

Lang, P. J. (1994). The varieties of emotional experience: A meditation on James-Lange theory. *Psychological Review*, *101*, 211–221.

Larson, J. T., To, Y. M., & Fireman, G. (2007). Children's understanding and experience of mixed emotions. *Psychological Science*, *18*, 186–191.

Moore, C., Harris, L., & Patriquin, M. (1993). Lexical and prosodic cues in the comprehension of relative certainty. *Journal of Child Language*, *20*, 153–167.

Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and décalage. *Developmental Psychobiology*, *40*, 255–265.

Morton, J. B., & Munakata, Y. (2009). Connectionist approaches to perseveration: Understanding universal and task-specific aspects of children's behavior. In J. P. Spencer, M. Thomas, & J. L. McClelland (Eds.), *Toward a unified theory of development: Connectionism and dynamic systems theory re-considered* (pp. 141–164). Oxford, UK: Oxford University Press.

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 6-year-olds' understanding of emotion in speech. *Child Development*, *74*, 1857–1868.

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

Solomon, D., & Yaeger, J. (1969). Effects of content and intonation on perceptions of verbal reinforcers. *Perceptual and Motor Skills*, *28*, 319–327.

Terwogt, M. M., & Rieffe, C. (2003). Stereotyped beliefs about desirability: Implications for characterizing the child's theory of mind. *New Ideas in Psychology*, *21*, 69–84.

Zelazo, P. D., & Frye, D. (1997). Cognitive complexity and control: A theory of the development of deliberate reasoning and intentional action. In M. Stamenov (Ed.), *Language structure, discourse, and the access to consciousness* (pp. 113–153). Amsterdam: John Benjamins.

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