

Title. Preschoolers' use of emotional prosody to resolve communicative ambiguity as a function of speaker conventionality

Introduction. Consider the following utterance: “*Look at my new haircut*”. This utterance can convey markedly different meanings if spoken in a sad-sounding voice versus a happy-sounding voice (e.g., Berman et al., 2010; Morton & Trehub, 2001). As illustrated by this example, understanding another’s communicative intent often involves the integration of the words spoken with non-linguistic cues. The focus of the current study is on one particularly influential cue that speakers can use to signal their intended meaning, namely emotional prosody. Emotional prosody is a paralinguistic cue that provides information about a speaker’s emotional state or disposition as expressed through variations in pitch level, pitch contours, and rate of speech (Banse & Scherer, 1996; Frick, 1985).

Between the ages of 3 and 5, children become increasingly attuned to a speaker’s emotional prosody and quickly integrate this information to guide language processing, such as when resolving ambiguous statements (i.e., statements where the meaning is unclear; Berman et al., 2010; Berman et al., 2013; Khu et al., 2018; San Juan et al., 2017). What continues to be debated, however, is the nature of the core mechanisms that underlie this ability and its development. For example, do children simply make statistical connections between emotional prosody and certain events or states-of-affairs (i.e., an associative account; Erickson & Thiessen, 2015; Perruchet, 2005; Perruchet & Pacton, 2005) or do they rely on active reasoning regarding how context, perceptual input, and mental states infer communicative intent (i.e., a socio-cognitive account; Goodman & Stuhlmuller, 2013)?

Although vocal emotions are universal and can be reliably identified across cultures (e.g., Bryant & Barrett, 2008), there is significant variability within and between people when it comes to emotion expression. As a result, emotional prosody serves as an interesting avenue of delineating the mechanisms that underlie children’s language comprehension. Recent research has shown that 4- and 5-year-old children suspend their use of emotional prosody to resolve communicative ambiguity when they are exposed to a speaker who is atypical in their use of emotional prosody (Thacker, et al., under review). This pattern, wherein a speaker who violates communicative norms will lead a listener to disregard emotional prosody cues, is highly compatible with a socio-cognitive account. What remains unclear, however, is how flexibly children can use emotional prosody to infer intent in response to a single speaker who uses emotional prosody in conventional and unconventional ways within the same interaction. The purpose of the current study was to evaluate the mechanisms that underlie 4- and 5-year-olds’ use of emotional prosody in resolving communicative ambiguity as a response to moment-to-moment changes in speaker conventionality.

Method. English-speaking 4-year-olds ($N=37$) and 5-year-olds ($N=29$) participated in a task that measured their implicit (i.e., looking) and explicit (i.e., pointing) sensitivity to emotional prosody during real-time language processing. Children were introduced to a speaker who used emotional prosody in a conventional manner (e.g., “*My soccer team won the championship*” in a happy voice) and an unconventional manner (i.e., “*My soccer team won the championship*” in a sad voice; see *Figure 1*). Using a mixed-factor design, participants were randomly assigned to one of two possible conditions that differed in terms of the order that the conventional or unconventional speaker was presented (i.e., Conventional Prosody Use 1st and Unconventional Prosody Use 1st). During test trials, participants were prompted to “*Look, look at X! Point to the X*” in either positive (i.e., happy) or negative (i.e., sad) emotional prosody, while being presented with 2 images that belonged to the same category but varied in terms of their

likelihood to be associated with positive (e.g., blooming flower) or negative (e.g., wilted flower) emotional prosody.

Results. Regarding children’s looking patterns, a linear-mixed effects model analysis was conducted using the fixed effects of age, condition, speaker type, and emotional prosody, as well as the random effects of subject and item. The analyses indicated a main effect of emotional prosody ($\beta = .69$, $SE = .30$, $t = 2.33$, $p = .03$; see *Figure 2*), where children were more likely to look at the target object during negative emotional prosody trials ($M = .63$, $SD = .48$) compared to positive emotional prosody trials ($M = .43$, $SD = .49$). Regarding children’s pointing patterns, a mixed-effects logistic regression was conducted using the same fixed effects and random effects. The analyses indicated a main effect of condition ($\beta = .24$, $SE = .11$, $t = 2.13$, $p = .03$), such that children were 1.28 times more likely ($B = 1.28$, 95% CI [1.02, 1.60]) to point to the target object in the Conventional Prosody Use 1st condition ($M = .60$, $SD = .12$) relative to the Unconventional Prosody Use 1st condition ($M = .54$, $SD = .14$).

Conclusions. The findings indicate that when 4 and 5-year-olds interact with a speaker who is both conventional and unconventional in their expression of vocal emotion, their use of emotional prosody is not uniformly disrupted. Children’s eye patterns demonstrated a negativity bias, such that they demonstrated more accurate looking in response to negative stimuli. Children’s pointing patterns demonstrated that when the speaker used emotional prosody in a conventional manner first, they continued to subsequent emotional prosody cues to resolve ambiguity. This effect was reduced when the speaker used emotional prosody in an unconventional manner first. This work demonstrates the flexible nature of children’s language comprehension patterns and aids in clarifying the pragmatic thresholds that 4- and 5-year-olds apply in their in-the-moment reasoning about a speaker’s communicative intent.


Visual Stimuli	Auditory Stimuli	Speaker Conventinality	Emotional Prosody
	<i>“My soccer team won the championship.”</i>	Conventional Unconventional	Positive (happy) Negative (sad)

Figure 1. Example of a teaching trial.

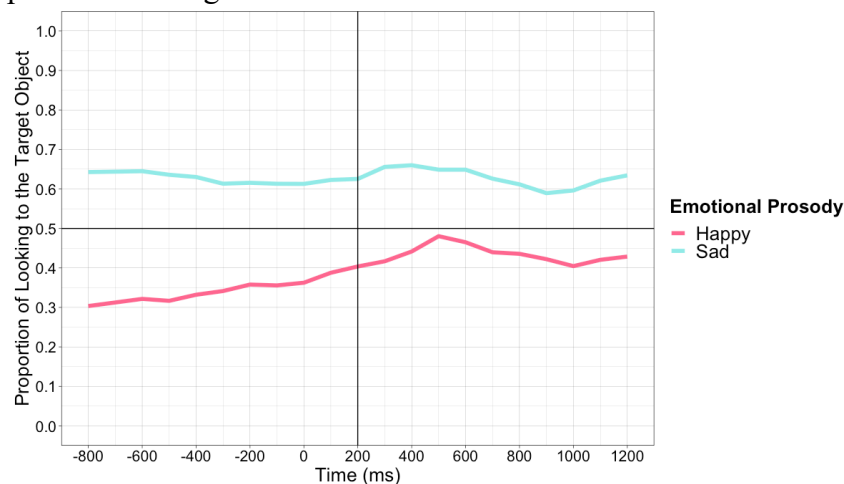


Figure 2. The proportion of looking to the target object (i.e., the object that matches the prosody of the speaker) as a function of the emotional prosody of the speaker; noun interval at 200ms.