Influences of Semantic and Prosodic Cues on Word Repetition and Categorization in Autism

Leher Singh and MariLouise S. Harrow

Purpose: To investigate sensitivity to prosodic and semantic cues to emotion in individuals with high-functioning autism (HFA).

Method: Emotional prosody and semantics were independently manipulated to assess the relative influence of prosody versus semantics on speech processing. A sample of 10-year-old typically developing children (n = 10) and children with HFA (n = 10) were asked to repeat words that were either emotionally congruent or incongruent in form and content (Experiment 1A). In a second task (Experiment 1B), the same participants were asked to classify stimuli on the basis of emotional prosody. A final experiment (Experiment 2) focused on sensitivity to congruence in a non-emotional source of variation: talker gender.

Results: The results revealed a selective impairment in spontaneous integration of prosodic and semantic cues to emotion in HFA; however, the same participants were able to categorize emotions on the basis of prosody under reduced task demands. Individuals with HFA were highly sensitive to another surface characteristic in speech: talker gender.

Conclusions: The study reveals impairment in the spontaneous integration of prosodic and semantic cues to emotion in HFA; however, insensitivity to surface detail, such as prosody, in HFA appears to be highly task dependent and selective to the domain of emotion.

Key Words: autism, prosody, development

Autism is a neurodevelopmental disorder associated with two primary areas of impairment: (a) deficits in social communication and interaction and (b) orientation toward repetitive and restrictive behaviors (American Psychiatric Association, 2013). The level of functioning varies considerably along the autism spectrum, with individuals with high-functioning autism (HFA) often possessing large vocabularies, using sophisticated sentence constructions and demonstrating average to above-average intelligence (Frith, 1991). However, these individuals commonly experience singular difficulty in the pragmatics of language (McCann & Peppe, 2003; Paul, Augustyn, Klin, & Volkmar, 2005; Paul, Orlovski, Marchinko, & Volkmar, 2009).

Pragmatics refers to the ability to integrate language into its appropriate context and to conform to standard communicative conventions that guide conversations (Pijnacker, Hagoort, Buitelaar, Teunisse, & Geurts, 2009). A facility with pragmatics enables a speaker to integrate context into conversation. Although pragmatically relevant cues are often not explicitly encoded in conversation, these details affect the semantic interpretation of speech and, as such, can inform the linguistic message (Levinson, 1983). An important carrier of pragmatic information is prosody: the rhythm and intonation of speech that is primarily driven by a constellation of temporal and intonational patterns in speech, one of which is vocal pitch. Even though some dimensions of prosody (e.g., emotional prosody) are defined as paralinguistic, speakers frequently vary the prosody of their voice for a linguistic purpose. There are several ways in which the perception of prosody informs the linguistic message on the part of listeners, such as the use of prosody to arrive at syntactic constituents (Snedeker & Trueswell, 2002), to distinguish questions from statements (van Heuven & Haan, 2002), to differentiate the first mention of a word from previous uses (Bock & Mazzella, 1983), and importantly, to impart vocal emotion (Scherer, 1986).

Although emotional prosody is considered to be a source of nonlinguistic variation, it strongly influences the linguistic message. For example, if a reader were to produce the sentence “Wow. This manuscript is absolutely
thrilling" in an emotionally flat or bored tone, the prosodic contour of his or her voice would contravene the literal meaning of the word *thrilling* such that the meaning of the sentence becomes drastically different from its literal form. In equal measure, concordances between word choice and emotional prosody (e.g., “Wow! This manuscript is absolutely thrilling!” spoken with pitch and amplitude variation associated with a state of excitement) can add emphasis, authenticity, or conviction to a speaker’s message. The objective of the current study was to investigate sensitivity to prosodic cues to emotion in relation to lexical cues to emotion in individuals with autism and typically developing (TD) individuals. A second objective was to investigate whether individuals with autism respond to surface cues to emotion in a manner that is distinct to other surface cues, such as those linked to a speaker’s identity. The purpose of this was to determine whether responses to surface variation in individuals with autism are specific to the domain of emotion or whether differences observed in autism extend to other forms of surface variation.

The issue of whether individuals with autism are typical in their perception and/or production of prosody has garnered much scientific interest. Atypical prosody in production was originally cited as a defining feature of autism by Kanner (1943) in his introduction of autism to the scientific community. Kanner’s initial observation has since been subjected to considerable empirical scrutiny (McCann & Peppe, 2003; Paul et al., 2005; Peppe, McCann, Gibbon, O’Hare, & Rutherford, 2007; Rutherford, Baron-Cohen, & Wheelwright, 2002). Empirical studies on prosody in speech production have revealed atypical prosodic production in autism in narrative production and structured contexts (Bonneh, Levanon, Dean-Pardo, Lossos, & Adini, 2011; Diehl, Watson, Bennetto, McDonough, & Gunlogson, 2009; Nadig & Shaw, 2012) as well as a reduced capacity for prosodic imitation in autism (Diehl & Paul, 2012). Differences in productive prosody have been thought to be a highly consistent feature of autism, with evidence from long-term studies that this is one feature that persists through development (Kanner, 1971; Ornitz & Rivto, 1976).

Individuals with HFA also demonstrate atypical perception of prosody under certain circumstances. For example, when examined on the ability to receptively identify single words as having either positive or negative affect based on prosody, children with HFA show significant impairment compared to their TD counterparts (Hesling et al., 2010; Jarvis-Pasley, Peppe, King-Smith, & Heaton, 2008; McCann, Peppe, Gibbon, O’Hare, & Rutherford, 2007; Peppe et al., 2007). In a more demanding paradigm, when asked to classify speakers’ emotional prosody according to a complex categorization system (e.g., concerned, sincere, regretful, insulted), individuals with HFA have been shown to perform more poorly than TD controls (Golan, Baron-Cohen, Hill, & Rutherford, 2007; Rutherford et al., 2002).

Similarly, when presented with verbal stimuli consisting of random strings of numbers spoken with an emotional tone, adults with HFA are significantly poorer than TD controls in categorizing stimuli into a prescribed set of five emotions (Philip et al., 2010). There is therefore a groundswell of evidence that individuals with HFA, though often fluent conversationalists, may have a specific deficit in the perception of emotional prosody.

At the same time, several studies have demonstrated comparable abilities in individuals with HFA and TD controls (Baker, Montgomery, & Abramson, 2010; Brennand, Scheppman, & Rodway, 2011; Grossman, Bemis, Plesa Skwerer, & Tager-Flusberg, 2010; O’Connor, 2007), suggesting that the perception of emotional prosody is not significantly impaired in individuals with HFA. Although it remains unclear why research in this area has yielded such inconsistent findings, possible explanations that have been ventured include effects of task demands, influences of verbal and nonverbal IQ on emotion perception, and heterogeneity within the individuals with HFA. With regard to task demands, studies of tasks involving more complex emotions; low-intensity emotions; and/or more complex tasks, such as face–voice matching, have reported group differences with greater consistency (e.g., Golan et al., 2007; Grossman & Tager-Flusberg, 2012; Loveland et al., 1995; Mazefsky & Oswald, 2007; Rutherford et al., 2002). Furthermore, individual participant characteristics, such as verbal and nonverbal IQ, have been shown to influence performance in emotion perception (see, e.g., Loveland et al., 1995, and Jones et al., 2011). This suggests that there may be multiple covariates within samples of individuals with autism that have not always been identified and statistically controlled, resulting in an inconsistent set of findings. In addition to participant-internal variation, varying methodological practices (e.g., full-spectral vs. filtered speech, unimodal vs. bimodal presentation) may have obscured consistent effects of affective prosody on emotion recognition, making it difficult to draw unequivocal conclusions about deficits in processing emotional prosody in individuals with HFA. Methodological variations are crucial to interpretation, given that individuals’ responses to speech are often highly task dependent. For example, prior research has demonstrated that unimodal versus bimodal presentation yields important differences in an individuals’ speech perception, with reduced face–voice integration associated with individuals with HFA as compared with TD individuals (de Gelder, Vroomen, & van der Heide, 1991; Smith & Bennetto, 2007).

When considering how tasks used in previous studies correspond to the demands of emotion recognition in natural speech, there are some noteworthy differences. Previous studies investigating sensitivity to emotional prosody in autism have largely used categorization paradigms in which participants are asked to label semantically neutral words varying in emotional prosody using a prescribed set of labels (Brennand et al., 2011; Golan, Baron-Cohen, Hill, & Golan, 2006; Grossman et al., 2010; Hesling et al., 2010; Jones et al., 2011; Lindner & Rosen, 2006; Loveland et al., 1995; McCann et al., 2007; Paul et al., 2005; Peppe et al., 2007; Philip et al., 2010; Rutherford et al., 2002). Semantic categorization tasks such as these are thought to be the product of postlexical decision making (Bradley &
Forster, 1987; Pexman, Hargreaves, Siakaluk, Bodner, & Pope, 2008). Traditionally, postlexical processes have been thought to be relatively slow and under conscious control given that the product of lexical processing must be explicitly represented in order for semantic categorization to take place (Fodor, 1983; Forster, 1981). Postlexical processing per se is often necessary in order to integrate linguistic information with contextual information to arrive at the correct semantic interpretation (Coulson & Federmeier, in press). However, the kind of postlexical processes engaged in paradigms typically used in previous studies of autism require participants to reason about emotional prosody in a structured fashion. More specifically, the act of assigning experimental stimuli to labels prescribed by the experimenter entails classifying emotional prosody according to a structured fashion. More specifically, the act of assigning experimental stimuli to labels prescribed by the experimenter entails classifying emotional prosody according to an externally specified taxonomy that may or may not correspond to the participant’s internal representation of emotional categories. As such, the behavioral demands of the task may not correspond directly to core processes mandated by everyday speech processing. By contrast, measures such as word recognition or word repetition are thought to represent online, automatic responses that tap the mental lexicon during the course of lexical processing (Coulson & Federmeier, in press). Such responses are thought to be faster, less likely to operate under conscious awareness, and unlikely to reflect the implementation of a strategy (Fodor, 1983; Luce & Pisoni, 1998; Swinney, 1979). Therefore, categorization paradigms commonly applied to studies of emotional prosody in autism may reflect activation at a particular stage of word processing. The specific types of categorization paradigms typically used in studies of autism, such as stimulus classification, have been shown to reflect higher levels of analysis than those implicated in basic word processing (Coulson & Federmeier, in press; Klauer, Musch, & Eder, 2005; Pexman et al., 2008).

Emotion recognition in natural discourse is arguably more complex than slotting words and phrases into a set of prescribed categories, but rather it is incumbent on the listener to correctly impute emotions to a speaker without an externally specified system. In natural discourse, there is free variation in emotional prosody (Pittman & Scherer, 1993), and efficiently recognizing the correct emotion intended by a speaker presents a challenge that may be underestimated by an explicit categorization paradigm. By examining implicit processing of emotional prosody it may be possible to engage more autonomous cognitive processes and to access participants’ internal emotion categories than those provided within the experiment. As such, these types of tasks may serve as a useful complement to the kinds of categorization paradigms that have prevailed thus far. For example, it is possible that the same stimulus set, varying in emotional prosody, may yield a different set of findings depending on the type and level of processing cued by the task.

Finally, in previous studies of emotional prosody perception in autism, semantic content is typically neutralized (or even filtered out) such that the sole focus of participants’ attention is intentionally directed toward emotional prosody. Although one can appreciate the methodological advantage of isolating effects of a particular independent variable, evidence from TD individuals suggests that prosodic and lexical features are processed independently and interdependently at a relatively early stage of lexical processing (see, e.g., Cutler, Andics, & Fang, 2011; Mullennix & Pisoni, 1990; Nygaard & Queen, 2008). This is perhaps to be expected given that the reconciliation of prosodic and lexical features has functional significance for language comprehension. In examinations of whether prosodic and lexical cues are processed interdependently or independently, the immediacy and automaticity with which participants avail themselves of both types of cues has become apparent: A classic finding from these paradigms is that participants attend to the form of a word even when they are directed to attend only to lexical features (see, e.g., Cutler et al., 2011; Mullennix & Pisoni, 1990). Therefore, using tasks that train the spotlight on prosody alone may not reflect the interdependent nature of prosodic and lexical processing.

A paradigm that has helped psycholinguists to determine how listeners reconcile prosody and semantic cues to emotion via online automatic processing is the congruence paradigm, previously used in typically functioning adults (Nygaard & Queen, 2008). In this task, emotional words (e.g., “cheer”) were spoken in a congruent (i.e., a happy tone of voice) and incongruent (e.g., a sad tone of voice) emotion, and participants were asked to simply repeat each word that they hear. The dependent measure was the speed with which they repeated the words (repetition latency). Congruence effects were observed: Adults were faster to repeat the word if the emotional prosody of the word matched its semantic content (e.g., happy words spoken in happy voices) than if there was a mismatch between prosodic and semantic cues to emotion (Nygaard & Queen, 2008). The dependent variables generated from congruence tasks—repetition latencies—are considered to be a very pure and implicit measure of lexical access (Forster, 1981) and are thought to engage the early stages of lexical access (Wurm, Vakoch, & Seaman, 2004), although it should be noted that word repetition may not reflect natural language processing associated with conversational discourse. Moreover, response latencies obtained from congruence tasks constitute a sensitive and easily accessible measure of word processing that can be potentially used with individuals with HFA without the task requirements (word repetition) challenging one population disproportionately. As such, congruence tasks provide a potential opportunity to investigate emotional prosody perception in autism in a novel and potentially informative way.

Sensitivity to congruence is an important component of social communication. The presence of congruence or

---

1Although a strict distinction between lexical and postlexical processing is not uncontested, there is general consensus that the relative emphasis on lexical/postlexical processes varies depending on the task at hand, with semantic categorization paradigms primarily reflecting postlexical processing and verbal repetition primarily reflecting early stage processing.
incongruence between prosody and word choice can alter the linguistic message, a property of language to which listeners appear highly sensitive. There is increasing evidence that adults automatically integrate emotional prosody and word choice, using both sources of information to arrive at word meaning (see, e.g., Ishii, Reyes, & Kitayama, 2003; Kitayama & Ishii, 2002). Moreover, a disconnect, or incongruence, between emotional prosody and word identity influences lexical processing (Wurm, Vakoch, Strasser, Calin-Jageman, & Ross, 2001). Studies investigating the integration of emotional prosody and word choice in typical adults have typically used online speech-processing tasks (e.g., Stroop task, lexical decision) and have converged on the conclusion that adults integrate the congruence or incongruence of emotional prosody and word choice into their semantic interpretation of speech in a way that is early (see, e.g., Kotz & Paulmann, 2007), automatic, and rapid. In light of this, one might expect an insensitivity to the relationship between emotional prosody and word choice to compromise a listener’s interpretation of the linguistic message.

The purpose of the current study was to investigate influences of prosody and semantic content on word processing using two different tasks—a congruence paradigm and semantic categorization—using the same stimulus set within the same set of individuals. In Experiment 1A, individuals with HFA and TD children were assessed on the efficiency with which they processed emotional stimuli that conflicted or concurred in prosodic and lexical cues to emotion using a congruence paradigm. After Experiment 1A, all participants were asked to categorize the same set of emotionally toned stimuli (Experiment 1B). We hypothesized that individuals with HFA would show selective deficits in integrating emotional prosody and word choice unlike their TD peers but that this pattern of results would be task dependent. More specifically, we hypothesized that verbal repetition would evince a reduced capacity for integrating emotional prosody and word choice in individuals with HFA but not in TD peers. By contrast, we hypothesized that, when asked to categorize the same stimuli according to a prescribed taxonomy, group differences between individuals with HFA and TD children may be neutralized. In a final experiment (Experiment 2), participants were administered a congruence paradigm with non-emotional stimuli. The purpose of this experiment was to qualify our results in integrating emotional prosody and word choice in individuals with HFA and a TD group.

Method

Participants. Ten children diagnosed with HFA and 10 TD children participated in this study. Children were matched on chronological age and formal language abilities across groups. Children with HFA had received the HFA diagnosis from a child psychiatrist and were referred for the study by the practitioner who made the diagnosis. All children received a diagnosis on the basis of behavioral observation and scores on Module 3 of the Autism Diagnostic Observation Schedule (Lord et al., 1989). TD children reported no known speech or language impairments and no relevant medical or psychological conditions and were healthy, well-functioning children performing at grade level. Participants’ ages ranged from 8:1 (years:months) to 13:0 for the participants with HFA, with a mean age of 10:7. For the TD group, age ranged from 8:8 to 12:0, with a mean age of 10:6. All children came from monolingual English-speaking homes. Both groups consisted of nine males and one female. Groups were matched on language skills (results of standardized testing are discussed later), with the exception of pragmatic skills, and were individually matched on chronological age.

All participants were given a suite of standardized tests to ensure that the groups were indeed comparable on formal language skills (vocabulary, sentence-level grammar, and morphology). Prior to participating in the experiments, all participants completed the Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF–4; Semel, Wiig, & Secord, 2003), a test of global language skills with several components including expressive language, receptive language, phonological skills and a Pragmatics Profile. In addition, two vocabulary tests were administered: (a) the Peabody Picture Vocabulary Test—Revised (Dunn & Dunn, 1981), to measure receptive vocabulary size, and (b) the Expressive Vocabulary Test (Williams, 1997), to measure productive vocabulary size.

Each participant’s parent completed a short questionnaire related to pragmatic and social development: the Social Communication Questionnaire (SCQ; Berument, Rutter, Lord, Pickles, & Bailey, 1999), which was developed to accompany the Autism Diagnostic Interview—Revised (ADI–R; Lord, Rutter, & Le Couteur, 1994), the predominant diagnostic instrument for autism spectrum disorders (ASDs). The SCQ is a parent checklist based on the ADI–R and is designed to evaluate communication skills and social functioning in children with autism. It can be administered to both children with autism and TD children. Scores on this test correlate with the severity of autism symptoms such that higher scores reflect a greater degree of pragmatic impairment (Chandler et al., 2007). The SCQ has been shown to be as effective as the ADI–R at discriminating participants with autism and TD individuals (Berument et al., 1999). Examples of items on the SCQ include questions about the child’s use of appropriate
facial expressions, the child’s interest in other children, and the child’s attentiveness to other people in his or her environment.

The results of standardized testing are displayed in Table 1 (standard scores). Comparing groups on their language abilities, there were similarities and differences across groups. In terms of productive and receptive vocabularies, the groups were highly similar, and Expressive Vocabulary Test and Peabody Picture Vocabulary Test scores did not differ across groups, $t(18) = -2.7, ns$, and $t(18) = -0.94, ns$, respectively. However, the individuals with HFA performed less well on the CELF-4 (only when the Pragmatics Profile was included), $t(18) = -3.23, p < .05$, compared with the TD group. In addition, the individuals with HFA were associated with greater impairment in social communication as evidenced by higher scores on the SCQ (i.e., greater pragmatic impairment) compared with the TD group, $t(18) = 7.87, p < .001$.

**Stimuli.** Stimuli consisted of 24 tokens digitally recorded by a female native English speaker and presented in two distinct prosodies (happy and sad). Stimuli consisted of emotional words (see Appendix A) varying in prosody. Emotional prosody was either incongruent with the lexical content of the word (e.g., happy words such as cheerful produced in a sad voice) or congruent with the lexical content of the word (e.g., cheerful produced in a happy voice). Words were chosen on the basis of being highly frequent and familiar to all participants. In their study of congruence effects with emotional stimuli, Nygaard and Queen (2008) also included neutral words and neutral emotional tones in their study. However, neutral stimuli were surprisingly not associated with congruence effects (i.e., neutral words spoken in neutral tones were associated with longer latencies than neutral words spoken in happy or sad tones), and neutral words spoken in happy or sad tones were associated with longer latencies than happy words spoken in sad tones and sad words spoken in happy tones, reflecting an anomalous pattern of results for neutral stimuli. This supports the argument that the time scale according to which neutral stimuli are processed may not be directly comparable to that of emotional stimuli, which are reportedly processed more rapidly than neutral stimuli (Kitayama & Howard, 1994; Niedenthal, Halberstadt, & Setterlund, 1997) and possibly by distinct neural mechanisms (Adolphs, 2002). For these reasons, neutral stimuli may be processed in comparable ways to positive and negative stimuli and were therefore excluded from the current study.

We implemented two stimulus validation procedures: (a) affective ratings and (b) acoustic analyses. To obtain affective ratings, 10 adults were asked to classify the stimuli into “happy” or “unhappy” categories. Adults correctly categorized the stimuli with 100% accuracy, grouping words produced in an emotionally positive tone of voice as “happy” and those produced in a neutral tone of voice as “unhappy” regardless of lexical content. Acoustic analyses were conducted on all stimuli to ensure that the speaker’s productions had a consistently positive or negative form independent of the content. The primary distinguishing features of “happy” and “sad” prosody are considered to be mean fundamental frequency and fundamental frequency range (Banse & Scherer, 1996), with happiness associated with higher fundamental frequency mean and range than sadness (although sadness is also associated with longer duration; van Bezooijen, 1984). Both measures were calculated for all words (see Figure 1). A 2 × 2 analysis of variance (ANOVA) with lexical valence (positive word/negative word) and prosodic valence (positive prosody/negative prosody) was conducted to determine effects and interaction of these factors on mean fundamental frequency. There was no main effect of lexical valence, but there was a main effect of prosodic valence, $F(1, 23) = 17.25, p < .0001$, with happy words having a higher fundamental frequency than sad words. There was no interaction of prosodic and lexical valence, $F(1, 23) = 0.16, ns$. With regard to fundamental frequency range, there was a main effect of prosodic valence on fundamental frequency range, $F(1, 23) = 58.4, p < .0001$, with happy words being associated with a higher fundamental frequency range than sad words. There was no main effect of lexical valence and no interaction of prosodic and lexical valence. The results of these analyses confirm that the speaker distinguished the primary carriers of happy and sad prosody (fundamental frequency mean and range) in a way that did not interact with the lexical content of the words.

**Procedure.** Each participant heard a set of 12 words in two prosodic patterns via computer speakers at conversational

<table>
<thead>
<tr>
<th>Standardized test (standard scores)</th>
<th>High-functioning autism</th>
<th>Typically developing</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVT</td>
<td>$M = 111, SD = 13.8$</td>
<td>$M = 114, SD = 14.5$</td>
</tr>
<tr>
<td>PPVT</td>
<td>$M = 115, SD = 8.3$</td>
<td>$M = 120, SD = 12.2$</td>
</tr>
<tr>
<td>SCQ</td>
<td>$M = 18.6, SD = 6.5$</td>
<td>$M = 2.3, SD = 1.2$</td>
</tr>
<tr>
<td>CELF-4</td>
<td>$M = 107.4, SD = 11.0$</td>
<td>$M = 117.5, SD = 5.31$</td>
</tr>
<tr>
<td>CELF-4 Pragmatics Profile</td>
<td>$M = 128.7, SD = 13.7$</td>
<td>$M = 182, SD = 14.8$</td>
</tr>
</tbody>
</table>

Note. EVT = Expressive Vocabulary Test; PPVT = Peabody Picture Vocabulary Test; SCQ = Social Communication Questionnaire; CELF-4 = Clinical Evaluation of Language Fundamentals, Fourth Edition.
volume (75 db) using the Brown Lab Interactive Speech System speech presentation software (Mertus, 2002). Half of the words were emotionally positive in lexical content and half were emotionally negative, resulting in 24 tokens. All of the words had emotional lexical content, such as cheerful, upset, laugh, and cry. Each word was presented twice: once with affect consistent with its meaning (e.g., glad spoken in a happy voice) and once with inconsistent affect (e.g., sad spoken in a happy voice). Both lexical valence and prosodic valence were therefore independently manipulated within subjects. The order of presentation of words was randomized for each participant, and participants were asked to repeat each word as quickly and as accurately as possible without starting to speak before the stimulus had finished playing. Onset of response was automatically detected by the experimental program, which allowed for precise measurement of the amount of time from stimulus offset to the onset of the participants’ repetition of words. No participant produced an intervening word (e.g., “Um”) prior to responding. Participants were given a maximum of 5 s to respond. If they did not respond within 5 s, the next trial would begin. There were a total of 24 trials.

Oral instructions provided to participants were to repeat the word as quickly as possible but to be sure to repeat the word correctly. Participants were tested using noise-attenuating headphones in a quiet room. No participant expressed queries about the task; neither did they express observations or questions about stimulus prosody or word meanings, nor did any participant express confusion about the task requirements. Each participant was tested individually.

**Results**

The primary dependent measure was repetition latency, specifically, the amount of time between offset of the stimulus and onset of a correct repetition by the participant. All participants repeated all of the words correctly within 5 s of the stimulus offset. A 2 × 2 ANOVA was conducted to examine main effects of the within-subjects factor (congruent/incongruent stimuli) and the between-subjects factor (HFA/TD) as well as possible interactions of these factors on repetition latencies. There was a main effect of congruence, $F(1, 18) = 18.1, p < .001, \eta^2 = .28$, with congruent stimuli named faster than incongruent stimuli. There was a main effect of group, $F(1, 18) = 24.68, p < .001, \eta^2 = .57$, with children with HFA showing slower response latencies overall than TD children. In addition, there was a significant interaction of group and congruence, $F(1, 18) = 27.85, p < .001, \eta^2 = .43$ (see Figure 2).

Planned comparisons revealed that TD children were significantly faster at repeating congruent words than incongruent words, $t(9) = –5.7, p < .001$, Cohen’s $d = –0.85$. By contrast, children with HFA did not show significant differences in repetition latencies to congruent versus incongruent stimuli, $t(9) = 0.9, ns$.

It is possible that congruence savings (this term refers to differences in latencies between congruent and incongruent stimuli and reflects the degree of facilitation afforded by form–content congruence) were less likely to be observed in the negative stimuli because there was a single emotion expressed prosodically (sadness) and three emotions expressed lexically (anger, worry/anxiety, and sadness). To ensure that congruence effects were not influenced by this, congruence savings were computed with three lexical items excluded (“angry,” “afraid,” and “mad”) from the negative stimulus set. When these items were excluded and a 2 × 2 (Congruence × Group) ANOVA was conducted on the remaining stimuli, there was a main effect of congruence, $F(1, 18) = 23.26, p < .001, \eta^2 = .46$, with congruent stimuli named faster than incongruent stimuli.

![Figure 1. Top panel: Mean fundamental frequency of stimuli (Experiments 1A and 1B). Bottom panel: Fundamental frequency range of stimuli (Experiments 1A and 1B). Error bars represent standard errors.](https://jslhr.pubs.asha.org/)

![Figure 2. Response latencies to word repetition for emotionally congruent and incongruent stimuli.](https://jslhr.pubs.asha.org/)
named faster than incongruent stimuli. There was a main effect of group, $F(1, 18) = 50.03, p < .0001, \eta^2 = .73$, with children with HFA showing slower response latencies overall than TD children. In addition, there was a significant interaction of group and congruence, $F(1, 18) = 8.63, p < .01, \eta^2 = .32$. Congruence savings were observed in the TD group, $t(9) = −5.09, p < .01$, Cohen’s $d = −1.81$, but not within participants with HFA, $t(9) = −1.45, ns$. The pattern of results was comparable to those emerging from the full stimulus set, consistent with acoustic analyses revealing a common set of correlates associated with worry, sadness, and anger (Banse & Scherer, 1996; Scherer, 1986).

Of further interest was whether congruence “savings” were associated with language skills and, in particular, with social communication abilities. Congruence savings were calculated by subtracting repetition latencies for congruent stimuli from those for incongruent stimuli to determine the extent to which congruence facilitates word processing within an individual. Congruence savings were correlated with the results of tests measuring social and pragmatic functioning. Scores on the SCQ were negatively correlated with congruence differentials, $r(20) = −.72, p < .01$ (two-tailed). By contrast, there were no significant correlations between congruence differentials and receptive or expressive vocabulary sizes or overall CELF–4 scores (see Table 2).

Overall, the results of Experiment 1A demonstrate that children with HFA and TD children demonstrate distinct responses when faced with congruent and incongruent cues to emotion. TD children demonstrated the expected interference associated with incongruence between prosodic and semantic cues to emotion as evidenced in typical adults (Nygaard & Queen, 2008). However, children with HFA were slower to repeat words in general than TD children, but they did not show congruence savings. A follow-up question that arose from this set of results is whether individuals with HFA were capable of detecting contrastive emotional prosody in this task or whether they were wholly dependent on lexical cues to emotion. In Experiment 1B, the same stimulus set was presented to both groups of participants but children were asked to classify the emotional prosody of the same stimulus set instead of simply repeating the words.

### Experiment 1B

The purpose of Experiment 1B was to determine whether children with HFA and TD children were differentially capable of labeling emotional stimuli on the basis of prosody. If children with HFA were simply not able to recognize the emotions conveyed through prosody in the stimulus set, this would suggest that their response patterns in Experiment 1A may have been attributable to difficulty detecting rather than integrating emotional prosody into word representations. We hypothesized that both groups of children would be comparable in their abilities to categorize emotional prosody within the stimulus set. The participant groups and stimuli were identical to those in Experiment 1A.

**Procedure**

Participants were presented with the same stimuli as those used in Experiment 1A. They were asked to decide whether the speaker sounded “happy” or “sad.” They were not given explicit instructions to ignore word meanings but were asked to listen only to the speaker’s tone of voice. Each stimulus item was presented once and the order of presentation was randomized. Stimuli were presented using the same stimulus presentation paradigm as in Experiment 1A (Brown Lab Interactive Speech Stimulation speech presentation software), and participants were given a 5-s response window. Responses were coded for accuracy and latency, and only first responses were considered. Accurate responses were those in which the participant correctly classified prosodically happy items as “happy” regardless of lexical content and prosodically sad items as “sad.” Participants’ responses were audio recorded, and accuracy scores (number of errors) were derived offline. In addition, response latencies were collected for correct responses only. Latency data were derived manually by measuring the time between stimulus offset and a first response. Voice keys were not used to derive data in this experiment because of a participant’s occasionally vocalizing prior to his or her response (e.g., “Um,” “Let’s see”). Participants were asked to express their responses orally.

#### Results

There were two dependent measures of interest: (a) accuracy (number of errors) and (b) response latencies for correct responses. The number of errors and response latencies are displayed in Figure 3. A $2 \times 2$ ANOVA was conducted to explore the effects of group (HFA/TD) and congruence (congruent/incongruent stimuli) on accuracy of responses. Results with the full stimulus set are presented first, followed by results with the words angry, mad, and afraid excluded (results presented in parentheses). Results revealed no main effect of group, $F(1, 18) = 0.39$, $ns$ ($F[1, 18] = 0.69, ns$) and a main effect of congruence.

<table>
<thead>
<tr>
<th>Standardized test</th>
<th>Experiment 1A: Emotion congruence savings (RT to congruent items minus RT to incongruent items)</th>
<th>Experiment 2: Gender congruence savings (RT to congruent items minus RT to incongruent items)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVT</td>
<td>$−0.9$</td>
<td>$.04$</td>
</tr>
<tr>
<td>PPVT</td>
<td>$.03$</td>
<td>$−.19$</td>
</tr>
<tr>
<td>SCQ</td>
<td>$−.72^*$</td>
<td>$−.59^*$</td>
</tr>
<tr>
<td>CELF–4</td>
<td>$.26$</td>
<td>$−.18$</td>
</tr>
</tbody>
</table>

*Note. RT = response time.*

$^*p < .01$. 

---

Table 2. Correlations between standardized testing and congruence savings (emotion and gender).
no main effect of group, $F(1, 18) = 1.49, ns (F[1, 18] = 1.99, ns). An ANOVA was also conducted with response latencies as the dependent variable. The results demonstrated no main effect of group, $F(1, 18) = 1.1, ns (F[1, 18] = 0.52, ns), and a main effect of congruence, $F(1, 18) = 9.56, p < .01, \eta^2 = .35$, with congruent stimuli labeled faster than incongruent stimuli. There was a significant interaction of group and congruence, $F(1, 18) = 8.87, p < .01, \eta^2 = .24, (F[1, 18] = 23.35, p < .001, \eta^2 = .32), with TD individuals showing significantly slower responses to incongruent than congruent items, $t(9) = 3.45, p < .01$, Cohen’s $d = –0.77), (t[9] = –10.01, p < .001, Cohen’s $d = –0.93). By contrast, children with HFA showed no significant differences in response latencies for congruent or incongruent responses, $t(9) = –0.12, ns (t[9] = –0.46, ns)$.

In the categorization task used in Experiment 1B, group differences were not apparent using the same stimulus set with accuracy as a dependent measure. Although one cannot directly compare tasks that vary in demands as well as in the dependent variable, the differences between Experiments 1A and 1B exemplify some of the challenges in drawing firm conclusions about word processing from a single task or measure. In summary, it appears that individuals with HFA are less sensitive to influences of emotional prosody in a word repetition paradigm but are equally accurate in classifying the same stimulus set in a categorization paradigm.

A natural question to arise is whether this response pattern on the part of individuals with HFA is specific to emotion. Many sources of surface variation exist in speech in addition to emotional prosody (e.g., talker gender, emphatic stress), broadly termed sources of indexical variation that have been shown to strongly influence speech processing across a variety of tasks in adults and children (Goldinger, 1996, 1998; McLennan & Luce, 2005; Nygaard, Sommers, & Pisoni, 1994; Singh, Morgan, & White, 2004). In addition, congruence effects have been observed using other domains of indexical variation. For example, congruence effects have been reported for talker age. Words with old associations (e.g., *shilling*) when produced by an older speaker were recognized faster and more accurately than when they were produced by a younger speaker (Walker & Hay, 2011). In addition, interference effects have been consistently demonstrated in studies with adults and children with incongruent speaker gender and semantic content. For example, when classifying stimuli on the basis of speaker gender, adults typically respond faster to congruent word and gender pairings, such as *man* produced in a male voice and *girl* produced in a female voice, compared with trials on which these pairings are realigned (Green & Barber, 1981, 1983; Most, Sorber, & Cunningham, 2007). Similar interference effects based on gender cues have been observed in children (Jerger, Martin & Pirozzolo, 1988; Most et al., 2007). Experiment 2 was designed to investigate whether such effects are observed in HFA in order to determine whether individuals with HFA are sensitive to form–content inconsistencies of speech outside of the domain of emotion.

**Experiment 2**

The purpose of Experiment 2 was to design and implement a study similar to Experiment 1A but with words varying in congruence along a non-emotional dimension that commonly varies in natural speech: talker gender. Experiment 2 aimed to investigate whether congruence effects transcend emotion in TD individuals and, if so, whether individuals with HFA are invulnerable to congruence effects in the domain of gender as well as in emotion. We hypothesized that individuals with HFA will respond comparably to TD peers when a non-emotional domain—in this case, gender—is manipulated. Participants were identical to those in Experiment 1A and 1B.

**Method**

**Stimuli.** Stimuli consisted of a set of 12 words (see Appendix B) with reference to male or female constructs. All stimuli were produced by a female and male native speaker of English, resulting in 24 tokens. Half of the words

---

**Figure 3.** Top panel: Accuracy levels (percentage correct) to the emotion categorization task. Bottom panel: Response latencies to the emotion categorization task.
had feminine referents, and half had masculine referents. Words could therefore be associated with gender lexically (e.g., husband is semantically associated with masculine gender) and indexically (e.g., any word spoken by a male is associated with masculine gender). Congruent stimuli referred to words with female referents produced by a female or those with male referents produced by a male. Incongruent stimuli referred to words with female referents produced by a male or words with male referents produced by a female. Stimuli were classified as being produced by a male or female by 10 adults with 100% accuracy.

Procedure. The stimulus presentation software and experimental paradigm were identical to those of Experiment 1A with the only distinction being the stimulus set. The order of experiments was the same for all participants (Experiment 1A, 1B, and 2).

Results

The primary dependent measure was repetition latency, specifically, the amount of time between offset of the stimulus item and onset of a correct repetition by the participant. All participants repeated all of the words correctly within 5 s of the stimulus offset. A $2 \times 2$ ANOVA was conducted to examine main effects of the within-subject factor (congruent/incongruent stimuli), and the between-subjects factor (HFA/TD), as well as possible interactions of these factors on repetition latencies. There was a main effect of congruence, $F(1, 18) = 59.46, p < .001$, with responses to congruent stimuli being significantly faster than those to incongruent stimuli. There was a main effect of group, $F(1, 18) = 9.53, p < .01$, with children with HFA showing slower response latencies overall than TD children. In addition, there was a significant interaction of group and congruence, $F(1, 18) = 10.96, p < .01$, with children with HFA showing slower response latencies than TD peers. In an effort to determine the bounds of congruence effects observed in Experiment 1A, form–content congruence was assessed in the domain of gender. Unlike in the case of emotion, both groups demonstrated expected congruence effects, with effects being greater for individuals with HFA compared with TD individuals. The size of gender congruence effects was inversely related to social communication skills within the sample, suggesting that autistic symptomatology was associated with a greater sensitivity to congruence based on cues to gender. By contrast, the size of congruence effects in the emotion task was directly related to social communication skills, suggesting that greater autistic symptomatology was associated with a reduced sensitivity

stimuli, $t(9) = -5.6, p < .001$, Cohen’s $d = -0.74$. The significant interaction can be explained by the finding that congruence savings (differences in response times to congruent and incongruent stimuli) was greater for individuals with HFA than TD individuals, $t(9) = 3.18, p < .05$, Cohen’s $d = 1.48$.

As in Experiment 1A, we were interested in whether congruence effects were associated with language skills and, in particular, with social communication abilities. Congruence effects were correlated with the results of standardized language testing. Scores on the SCQ were positively correlated with congruence savings, $r(20) = .59, p < .01$ (two-tailed), suggesting that individuals with poorer social communication abilities showed larger congruence effects. As in Experiment 1A, there were no significant correlations between congruence effects and receptive or expressive vocabulary sizes or overall CELF–4 scores (see Table 2).

### General Discussion

The purpose of the current studies was to compare children with HFA and TD peers on processing of prosodic and lexical cues to emotion in a congruence paradigm (Experiment 1A). We sought evidence of integral processing of prosodic and lexical cues by investigating repetition latencies to stimuli that were congruent in form and content compared with those that were incongruent. Congruence effects were investigated using stimuli varying in emotion (Experiment 1A) and those varying in gender (Experiment 2). First, in both tasks, individuals with HFA were slower to repeat items, suggesting a delay in basic psycholinguistic processes such as verbal repetition as compared with TD individuals. Second, in the emotional congruence tasks, participants with HFA did not show the expected emotional congruence effect demonstrated in TD peers and previously observed in typical adults (Nyggaard & Queen, 2008) and were not significantly different in their repetition latencies for congruent versus incongruent items. In Experiment 1B, both groups were equally accurate at classifying the same stimuli according to emotional prosody. In an effort to determine the bounds of congruence effects observed in Experiment 1A, form–content congruence was assessed in the domain of gender. Unlike in the case of emotion, both groups demonstrated expected congruence effects, with effects being greater for individuals with HFA compared with TD individuals. The size of gender congruence effects was inversely related to social communication skills within the sample, suggesting that autistic symptomatology was associated with a greater sensitivity to congruence based on cues to gender. By contrast, the size of congruence effects in the emotion task was directly related to social communication skills, suggesting that greater autistic symptomatology was associated with a reduced sensitivity

---

2Verbal repetition engages a sequence of processes including auditory word recognition, motor planning, and speech production. As such, in the event that overall latencies differ between clinical populations and TD peers, the precise locus of delay remains unidentifiable.
to congruence based on cues to emotion. In the aggregate, our results attest to an emotion-specific inability to spontaneously integrate form and content in individuals with HFA in an online speech-processing task in which prosody and semantics vary within a single word. Although a direct comparison of Experiments 1A and 1B are not appropriate given the differing task demands and dependent measures (repetition latencies vs. accuracy of classification), it appears that the same participants demonstrated different levels of facility with emotion processing depending on how the emotion processing task is structured. It should be noted that the absence of congruence effects in the labeling task may be due to more relaxed time constraints associated with categorization versus repetition.

A surprising finding was that analyses of individual differences in social communication skills as measured by the SCQ correlated divergently with emotion and gender congruence savings. On the one hand, the finding that emotion congruence savings would be directly related to social communicative skills is perhaps to be expected in light of the strong evidence basis for deficits in emotion perception in autism. Compromised processing of emotional stimuli is a widely documented trait of autism, and a negative association between the severity of autistic symptomatology and emotion recognition abilities has been reported in previous studies (e.g., Bal et al., 2010; Boraston, Blakemore, Chilvers, & Skuse, 2007; Humphreys, Minshew, Leonard, & Behrmann, 2007). On the other hand, although it may appear surprising that autistic symptomatology is associated with greater savings in gender congruence, there is a growing pool of evidence demonstrating that individuals with HFA outperform their TD peers in attentiveness to stimulus form even when form details are not central to the task (e.g., Jarrold, Gilchrist, & Bender, 2005; Joseph, Keehn, Connolly, Wolfe, & Horowitz, 2009; Plaisted, O’Riordan, & Baron-Cohen, 1998). Moreover, the extent of enhanced attentiveness toward stimulus form has been shown to correlate positively with autistic symptomatology (Joseph et al., 2009). These areas of anomalously high attentiveness—reported in both visual and language perception—have been advanced as candidate causes of some aspects of the autistic phenotype, such as atypical face processing and hyperlexia (Frith & Snowling, 1983; Samson, Mottron, Soulieres, & Zeffiro, 2011). Although TD individuals also encode stimulus features such as gender in the voice and process these features integrally with linguistic content (e.g., Mullennix & Pisoni, 1990), it appears, on the basis of the increased congruence savings evidenced in the children with HFA as compared with the TD group, that individuals with HFA depart from their TD peers in being extra-attentive to such details in speech.

One could interpret this as evidence for appropriately integrated representation of words in individuals with autism. However, in everyday speech processing placing undue weight on nonphonemic details such as talker gender presents a potential cost to language processing given that these details are typically secondary to word meaning in comparison to semantic cues (Mullennix & Pisoni, 1990). In sum, the divergences in association between congruence savings for gender and emotion tasks with SCQ scores may be due to two attested hallmarks of autistic symptomatology: (a) a selective deficit in emotion perception and (b) an acute sensitivity to stimulus form. The causal arrows connecting gender congruence savings, emotion congruence savings, and social-communicative skills cannot be drawn on the basis of the current findings. In addition, the sample size of the current study limits our interpretation of individual differences analyses in particular. Further research could venture greater specificity about how constructs herein relate to one another by measuring collateral abilities as possible mediators, such as attentiveness to form using nonlinguistic stimuli (e.g., visual search tasks) or integral processing of dual streams of information presented through different modalities (e.g., gesture–speech integration, as assessed by Silverman, Bennett, Campana, & Tanenhau, 2010).

A long-standing debate in autism is the extent to which the perception of mental states, including emotions, is selectively impaired (Baron-Cohen, 1989, 1991; Frith, 2003; Ozonoff, Pennington, & Rogers, 1990) compared with other types of events. In reality, very few studies have directly compared the processing of emotional—or, more broadly, mentalistic—stimuli with nonmentalistic stimuli within the same paradigm and modality. To the extent that we can compare results of Experiments 1A, 1B, and 2, the current study bears on this issue in demonstrating that emotional voice cues are marginally integrated in individuals with HFA, whereas non-emotional voices cues appear to be strongly represented in the speech percept. This points to dissociations in voice processing in emotional versus non-emotional sources of variation in individuals with HFA but not in TD peers. Furthermore, influences of emotional prosody were potent when accuracy was measured in a semantic categorization task but not when response times were measured in a word repetition task. This is consistent with previous evidence that individuals with HFA may show the capacity for perceiving emotions in the voice comparable to TD peers, but they may spontaneously integrate these cues less readily than TD controls (Begeer et al., 2008; Chevallier, Noveck, Happe, & Wilson, 2011). Therefore, TD children and those with HFA may differ significantly not in the basic ability to perceive contrastive emotions using prosodic cues but rather in the communicative relevance ascribed to emotional cues in speech perception. Relative to semantic cues, prosodic cues to emotion may be deemphasized in individuals with HFA, which may contribute to the widely reported deficits in abstracting nonliteral meaning in HFA from prosodic cues alone (e.g., McCann & Peppe, 2003; Peppe, 2009).

Analogous effects have been demonstrated in word learning tasks where individuals with autism are sensitive to the sound properties as well as semantic properties of words in equal measure to their TD peers but differ in their capacity to consolidate these sources of information into a unified lexical representation (Norbury, Griffiths, & Nation, 2010). Taken together, these findings tentatively suggest an atypical potential for form–content integration in online
speech processing tasks rather than atypical perception or attentiveness to emotional form, although such a hypothesis merits further targeted and systematic testing.

When placed within the context of previous studies investigating perception and processing in vocal affect in ASDs, the current study underscores the strong influences of experimental methodology and choice of dependent variable on participants’ performance. The current task incorporated a set of simple emotions into an online verbal repetition task (Experiment 1A) associated with early-stage lexical processing and a binary classification paradigm (Experiment 1B) associated with late-stage postlexical decision making. Although the stimulus set and participants were invariant across Experiments 1A and 1B, the difference in task demands elicited very different results, with both groups appearing quite different when repetition latencies were tracked and more similar when semantic categorization was used. This points to effects of emotional prosody on speech processing in ASDs being highly susceptible to effects of task demands. Therefore, it is perhaps to be expected that the studies thus far in this area have generated results that are seemingly difficult to reconcile with one another. Conflicting data also exist within the area of emotional recognition in the face in ASDs, with some studies demonstrating deficits in emotion recognition (see, e.g., Grossman & Tager-Flusberg, 2008; Philip et al., 2010) and other studies demonstrating little difference between individuals with ASDs and TD peers (see, e.g., Gepner, Deruelle, & Grynfeltt, 2001). A common set of confounds is sometimes raised to account for inconsistencies in emotion recognition in the face as well as in the voice across studies, such as IQ, severity of symptoms, and chronological age (e.g., Harms, Martin, & Wallace, 2010; Jones et al., 2011). Although it is undoubtedly important to identify confounds and to statistically define their effects, it is also possible that experimental paradigms (and their associated task demands) themselves constitute “effect modifiers” and can result in effects or non-effects of emotional variation on language processing. This possibility invites an integrative approach that combines different methodologies while identifying the putative psycholinguistic processes yoked to different paradigms over a mono-method approach that ascribes greater validity to one paradigm over another. Moreover, in the current study we incorporated a limited set of words in the interests of working within the limits of children’s attentiveness; it is likely that a larger (and semantically varied) set of words may allow for a more detailed analysis of the effects of word and emotion types. In addition to a limited number of words, the mapping between words and their intended emotions can vary in strength and nature of association. A more in-depth investigation that considers the effects of these mappings would advance our understanding of how particular form–content correspondences are processed in individuals with autism. Furthermore, vocal emotion is a highly complex construct comprising a multiplicity of interacting cues; as a consequence, the correlates of vocal emotion are often very difficult to reliably define at the level of individual emotions (Scherer, Johnstone, & Klumzeyer, 2003). It is possible that conceptualizing emotion as a single construct in an experimental paradigm is at odds with the inherent complexity of emotionally toned speech. Perhaps a greater emphasis on researching sensitivity to emotion “families” (e.g., high arousal vs. low arousal, hot emotions vs. cold emotions) would yield a clearer picture of how emotional subtypes are perceived and processed in individuals with ASDs.

Although our study reveals robust group differences between individuals with HFA and TD children, it must be noted that the limited sample size constrains our interpretations of results. First, there is greater scope for uncontrolled effects of individual idiosyncratic tendencies than would be expected with a larger sample. Such idiosyncrasies may not be a function of diagnostic category yet may influence the group mean quite significantly in a small sample. Second, in experimental studies it is often presumed that the “within” level of the independent variable (group) exists a fundamental homogeneity and that there is greater within-group similarity than between-group similarity. However, as with any sample, clinical or TD, such homogeneity is often difficult to define and confirm or disconfirm given that the degree of homogeneity is highly dependent on the particular constitution of each group. This issue is particularly relevant to small sample sizes. A larger, more extensive follow-up would help confirm the patterns of results revealed in this study.

The current study invites further research to inform our conclusions about the domain dependence of congruence effects in autism given that the issue of whether emotion processing is a specific area of impairment remains a significant theoretical controversy. The non-emotional independent variable herein, gender variation, is presumably categorically perceived, whereas emotion variation is multidimensional and continuously varying (Feldman, 1995; Pittman & Scherer, 1993). Further research could probe the extent to which an invulnerability to congruence in HFA is specific to emotion by manipulating other continuously varying sources of within-speaker variation in a similar paradigm, such as prosodic stress. Prosodic stress has been extensively investigated in HFA but primarily in the domain of production (Hubbard & Trauner, 2007; Paul, Bianchi, Augustyn, Klin, & Volkmar, 2008; Shriberg et al., 2001; Simmons & Baltaxe, 1975). Comparing effects of within-talker variation that are emotional or non-emotional in nature (e.g., affective prosody vs. prosodic stress) would shed further light on the extent to which processing of emotional prosody is selectively impaired in autism as compared to other within-talker sources of variation and may provide a stronger basis for direct comparison than gender and emotional prosody.

The current study contributes to our understanding of HFA by demonstrating a reduced sensitivity to prosodic and semantic congruence in emotional stimuli using a repetition task. Although findings from prior studies on emotion processing in HFA are somewhat hard to reconcile, the current study suggests that effects of emotional prosody variation on speech processing in HFA are highly
vulnerable to the experimental method. Performance accuracy on classification tasks that are traditionally linked to postlexical processes elicited the appearance of intact emotion processing in participants with HFA compared with TD peers. In contrast, a repetition task typically linked to online implicit processes elicited the appearance of a deficit in form–content integration in HFA that was specific to the domain of emotion. Sensitivity to form–content congruence was correlated with social communication skills, with stronger social communicative skills associated with reduced sensitivity to emotional congruence and increased sensitivity to non-emotional congruence (see Table 2). In addition to informing conclusions about emotion processing in HFA, the current study suggests that social communication abilities are not inconsequential for processing non-emotional stimuli in HFA; specifically, an overattentiveness to form–content congruence in non-emotional domains may relate to the strength of social communicative skills associated with autistic symptomatology. Finally, the current study suggests that the choice of dependent variable, task demands, and the target of attention may strongly influence the typicality of emotional processing observed in HFA and introduces a methodological complement to widely attested emotion-categorization paradigms.

Acknowledgments

We thank the participants and their families for their involvement in this study. We are grateful to Sarah Nestor for assistance with this research.

References


Appendix A
Stimulus Words Used in Experiments 1A and 1B

Awesome
Cheerful
Glad
Happy
Laugh
Smile
Afraid
Angry
Cry
Mad
Sad
Upset

Appendix B
Stimulus Words Used in Experiment 2

Girl
Female
Housewife
Lady
Missus
Woman
Boy
Husband
Guy
Man
Mister
Sir