LANGUAGE AND COGNITIVE PROCESSES
2007, 22 (2), 247–271

Psychology Press
Taylor & Francis Group

Receptive prosody in adolescents and adults with Williams syndrome

Daniela Plesa Skwerer
Boston University School of Medicine, Boston, MA, USA

Casey Schofield
Binghamton University, State University of New York, New York, NY, USA

Alyssa Verbalis
University of Connecticut, Storrs, CT, USA

Susan Faja
University of Washington, Seattle, WA, USA

Helen Tager-Flusberg
Boston University School of Medicine, Boston, MA, USA

People with Williams syndrome (WS) are known to use prosodic devices extensively in conversation and narratives, but their ability to interpret prosody to comprehend speakers' communicative intentions and emotional states has not been investigated systematically. We present findings from three experiments probing sensitivity to lexical stress and to affective prosodic cues

---

Correspondence should be addressed to Daniela Plesa Skwerer, Department of Anatomy and Neurobiology, Boston University School of Medicine, 715, Albany Street L-814, Boston, MA, 02118; E-mail: dplesas@bu.edu

This research was supported by a grant from the National Institute of Child Health and Human Development (RO1 HD 33470) and by grant M01-RR00533 from the General Clinical Research Center program of the National Center for Research Resources, National Institutes of Health to Boston University School of Medicine.

We thank Margaret Kjelgaard for providing the stimuli and design suggestions for Experiments 2 and 3. We express our sincere gratitude to the National Williams Syndrome Association and New England regional chapter for their help in recruiting participants; and to the families and individuals who participated in this study.

http://www.psypress.com/lcp DOI: 10.1080/01690960600632671

considered at several processing levels. Adolescents and adults with WS were compared with age, IQ and receptive vocabulary-matched participants with learning or intellectual disabilities (LID), and with age-matched normal controls (NC). The WS group performed significantly better than the LID group only in recognising emotional tone of voice in filtered speech. Results reflect a relative sensitivity in the WS group to *affective prosody*, while the ability to use *linguistic prosodic* cues for semantic interpretation remains constrained by perceptual and cognitive limitations, suggesting a possible dissociation in sensitivity to different types of prosody in this population.

Prosody, manifested as the variations in intonation, rhythm, stress, and melody of speech, serves a variety of functions in communication. It provides clues about the communicative intent and emotional states of the speaker, it conveys information regarding the syntactic structure or type of an utterance (e.g., questions, statements), and helps in disambiguating word meanings. Thus, different prosodic elements in speech play different roles, both linguistic and paralinguistic (Cutler, Dahan, & van Donselaar, 1997; Paul, Augustyn, Klin, & Volkmar, 2005). For the purposes of this paper we will distinguish between *linguistic* prosody, involved in communicating semantic and syntactic information, and *affective* prosody, involved in conveying emotions and affective states of the speaker, often enriching the expressiveness and social interactive aspect of communication.

Sensitivity to prosody in others' speech and the ability to use prosody appropriately to engage the listener, are essential components of a person's repertoire of social communicative skills. However, there has been relatively little systematic research examining the relations between receptive or expressive prosody and social competence, especially in people with learning or intellectual disabilities. Studies that have explored relations between receptive or expressive prosodic abilities and aspects of theory of mind in non-literal utterances have varied widely in methodology and have often reported conflicting findings both in typically developing children (Ackerman, 1983, 1986; Capelli, Nakagawa, & Madden, 1990; de Groot, Kaplan, Rosenblatt, Dews, & Winner, 1995; Dews et al., 1996; Keenan & Quigley, 1999; Winner & Leekam, 1991), and in people with autism (Rutherford, Baron-Cohen, & Wheelwright, 2002; Happé, 1993). Since prosody represents a key aspect of social communication, and because communicative dysfunction is one of the core symptoms of autism, it is not surprising that most research on prosodic abilities in people with developmental disorders has focused on individuals with autism spectrum disorders (see McCann & Peppe, 2003 for a review). In contrast, according to anecdotal reports, social communication appears to be a relative strength for people with WS (Jones et al., 2000), who are frequently described as very engaging in conversation and in narrative tasks. However, their ability to interpret and use prosody to comprehend others' communicative intentions and emotional states has not been examined systematically.

WS is a neurodevelopmental disorder caused by a microdeletion on the long arm of chromosome 7 (7q11.23), a region that includes about 20 genes. It is characterised by a unique phenotype that typically includes cardiovascular disease, connective tissue abnormalities, cranio-facial dysmorphology (referred to as *elfin facies*), and an unusual combination of cognitive and behavioural features (Morris & Mervis, 1999; Udwin & Yule, 1991). Initial small-scale studies of people with WS almost invariably noted what appeared to be a remarkable sparing in language abilities in the context of mental retardation and profound impairments in specific areas of cognition, especially visual-spatial-construction skills (Bellugi, Marks, Bihrle, & Sabo, 1988; Bellugi, Wang, & Jernigan, 1994). Although claims of 'intact' language have not stood up to closer systematic examination (Karmiloff-Smith et al., 1997), relative strengths in auditory memory (Mervis et al., 2000) and an undeniable impression of expressiveness in the spontaneous speech of people with WS have been consistently noted by those interacting with individuals with WS, in both naturalistic situations and during structured laboratory tasks (Jones et al., 2000).

Research on the behavioural and personality profile of people with WS has revealed a complex picture of sociability and empathy, coupled with poor social relationships and difficulties in social functioning (Gosh & Pankau, 1997; Klein-Tasman & Mervis, 2003). This paradoxical combination calls for careful investigation of the different sets of skills involved in communication and more generally, in social functioning in people in WS, by going beyond global evaluations of socio-communicative strengths and weaknesses, and examining both receptive and expressive abilities in varied contexts. So far studies of people with WS have focused mostly on expressive language in narrative tasks and in conversation.

One of the commonly used narrative tasks requires participants to tell a story based on a picture book. For example, Bellugi and her colleagues (Bellugi, Lichtenberger, Jones, Lai, & St George, 2000; Harrison, Reilly, & Klima, 1995; Losh, Bellugi, Reilly, & Anderson, 2000; Reilly, Klima, & Bellugi, 1990; Reilly, Losh, Bellugi, & Wulfeck, 2004) found that children and adolescents with WS used significantly more evaluative devices, such as character speech, sound effects, affective state words, emphatic markers and vocal prosody, compared with age-matched participants with Down syndrome and normal controls of the same mental age. Harrison et al. (1995) reported that adolescents and adults with WS used significantly more evaluative devices and affectively enriched language in spontaneous conversation collected in the course of a semi-structured biographical interview. However, a closer analysis of the functional role of the expressive devices used by individuals with WS in narratives and conversation suggests that they serve mainly as 'audience hookers' – as means of socially engaging the listener – while an analysis of the thematic and structural aspects of the narratives produced by children with WS revealed significant difficulties with the referential content and coherence of their stories, compared with both typically developing children and children with language impairment (Reilly et al., 2004). Thus, the use of expressive prosody appears to be a manifestation of the hypersociability characteristic of people with WS, which may or may not be appropriate for enhancing the quality of their communicative exchanges.

The set of experiments reported in this paper extends this line of research by focusing on *receptive prosody*. Two main questions guided this investigation of adolescents and adults with WS. First, is their strong motivation for social engagement through communicative means reflected in sensitivity to prosodic aspects in others' speech? Second, do people with WS appreciate the communicative role of different types of prosodic cues in conveying linguistic or affective information in speech?

We report on the performance of adolescents and adults with WS, compared with a control group of age-matched typically developing individuals (NC) and a group of individuals with learning or intellectual disability (LID) matched on age, IQ and vocabulary knowledge with the WS group, in three experiments. The first experiment examined the use of linguistic prosodic cues in interpreting the meaning of lexically ambiguous word-pairs. In the second experiment we assessed the ability to judge either the emotional prosody or the emotional semantic content of sentences heard in conditions of congruent and incongruent combinations of vocal prosody and verbal content. In the third experiment we probed sensitivity to affective prosody in low-pass filtered speech that eliminated access to the semantic content in spoken sentences.

Most previous research on lexical ambiguity processing has focused on the time course or on the sentential context influencing the selection of an appropriate meaning of ambiguous words, such as *hot dog* (see Nygaard & Lunders, 2002 for review). Although context is often critical for disambiguating lexical items, in spoken language potential ambiguities can be resolved by the prosodic pattern with which the words or sentences are uttered (Beach, 1991; Ferreira & Dell, 2000; Kjelgaard & Speer, 1999; Nespor & Vogel, 1987). For instance, the strong-weak stress pattern of two-syllable lexical structures can help differentiate the intended meanings of identical word-pairs, such as BLACKbird (an avian species) and black BIRD (any bird of black colour), or MAKEup (cosmetic item) and make UP (agreeing to end a quarrel). Developmental studies suggest that the ability to use stress to disambiguate word meaning emerges after age 10 in typically developing children (Vogel & Raimy, 2002).

Our goal in Experiment 1 was to investigate whether individuals with WS would be sensitive to prosody when it is the *only* source of information that can resolve the ambiguity of dual-meaning word-pairs. Prosody in this case plays exclusively a linguistic role. In this context we predicted that the individuals with WS would perform similarly to age, IQ and language matched peers with LID in using linguistic prosody as a clue to word meaning, and that both these groups would be less able than age-matched normal controls to use prosody for lexical disambiguation.

In spoken communication prosody also conveys affective and attitudinal states of the speaker, especially through intonational features of the voice (e.g., pitch variations and timing), as well as volume, speech rate, and voice quality. In most communicative situations prosody is congruent with and complements the content of speech (e.g., describing a happy event in a jubilant tone of voice). However, incongruent information in the two channels may be used deliberately to modify the meaning of the verbal content, for example to express irony or sarcasm. Prosodic information expressed in affective tone of voice can also be processed implicitly when people attend primarily to the content of speech, and may influence aspects of semantic processing.

Studies with normal adult responders in several cultures have shown the presence of an interference effect from information in the speaker's vocal tone on judgements of word or sentence meaning, and vice versa (Kitayama, 1996; Kitayama & Howard, 1994; Schirmer, Kotz, & Friederici, 2003). Moreover, cross-cultural experimental studies have shown that divergent practices of interpersonal communication, typical of different cultures, lead to reliable biases in the allocation of attentional resources to verbal content or to contextual cues, such as vocal tone in speech communication (Kitayama, 2002). Such differences in communicative style have been found between European/Western languages and cultures, in which verbal content serves as the primary means for conveying information, and Asian/Eastern ones, in which contextual information, including tone of voice, play a more prominent role. Experimental studies have found that American participants attended more to verbal content whereas Japanese participants attended more to vocal tone (Ishii, Reyes, & Kitayama, 2003; Kitayama & Ishii, 2002). These findings indicate that specific biases in processing information in the verbal and paralinguistic channels may be expected, depending on which aspects of the communicative exchange participants are particularly attuned to.

Although all the participants in the experiments reported in this paper were native speakers of English, we hypothesised that the individuals with WS might show a distinctive sensitivity for the speaker's vocal prosody compared with controls. To test this hypothesis, in Experiment 2 we presented two tasks that involve metalinguistic judgements: one task required an explicit judgement of the affective information (i.e., *happy, sad, neutral*) conveyed in prosody while disregarding the sentence content, whereas the second task required an explicit judgement of the emotional content of the sentence while ignoring the tone of voice of the speaker. Research has demonstrated that when a discrepancy between vocal prosody and verbal content is present, adults typically rely on affective prosodic cues to interpret the speaker's emotional state or communicative intent, while children younger than about 10 years focus more on verbal content, disregarding the affective prosodic cues, even when instructed to disregard sentence content (Morton & Trehub, 2001). This difficulty in attending to prosodic information in incongruent utterances has been linked to limitations in executive function abilities related to dealing with conflicting task demands (Morton, Trehub, & Zelazo, 2003), rather than to lack of sensitivity for paralinguistic emotional cues in speech (Borke, 1971; Fernald, 1992; Friend, 2000).

Our third experiment was designed to control for the potential executive demands of Experiment 2, by asking participants to judge affective prosody in the absence of comprehensible content information in speech. By obscuring the content information by low-pass filtering the speech stimuli, this experiment provided a purer measure of sensitivity to emotional paralanguage, thus we predicted that the participants with WS would perform better than IQ and vocabulary matched controls on this task.

EXPERIMENT 1: USE OF LINGUISTIC PROSODY IN LEXICAL AMBIGUITY RESOLUTION

This experiment was designed to test the ability to use prosodic stress cues in judging the meaning of ambiguous lexical constructs (word-pairs).

Method

Participants. Three groups of adolescents and adults, ranging in age from 12 to 35 years, participated in this experiment: 37 individuals with WS (22 females and 15 males), 36 individuals with learning or intellectual disabilities (LID; 21 females and 15 males) and 42 typically developing individuals (NC; 31 females and 11 males). All participants were native English speakers.

The participants with WS were recruited through the Williams Syndrome Association and all had their diagnosis of WS confirmed by genetic testing (i.e., the FISH test). The participants with LID were recruited through a residential school serving this population. The NC participants were recruited from local schools and universities. Participants were administered the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990) and the Peabody Picture Vocabulary Test-III (PPVT-III; Dunn & Dunn, 1997) to assess IQ and verbal knowledge. Table 1 presents the characteristics of the three groups. All three groups were matched on chronological age, $F(2, 112) = 1.49$, $p = .23$. The participants with WS were matched to the LID participants on chronological age, $p = .44$; IQ, $p = .97$; and PPVT standard scores, $p = .81$ (Scheffé test).

Stimulus materials. Audio stimuli were created using PRAAT audio recording software. A female amateur actress, whose native language was American English, recorded speech samples of 50 two-syllable word-pairs on a Dell Latitude computer. In the initial set of stimuli, 30 word-pairs had double meaning, and the remaining 20 were semantically unambiguous (foils). Each stimulus was recorded a minimum of three times in each of three conditions: (1) with the stress placed on the first syllable; (2) stress on the second syllable; (3) equal stress throughout. Three examples selected by the researchers from each of these conditions for each two-syllable construct were analysed for acoustic qualities using PRAAT speech analysis program. Acoustic measurements of average pitch (F_0), average amplitude/intensity and duration were taken for each syllable to ensure that the syllables differed significantly in their stress pattern. The following criteria (cf. Lieberman, 1960; Klatt, 1976) were used to select the best exemplars for the experimental test stimuli: the stressed syllable had to be pronounced with longer duration, higher pitch and greater intensity than the unstressed syllable. The acoustic measurements confirmed the expected differences between the stressed and unstressed monosyllabic words in the selected word-pairs. As expected, in the equal stress condition these differences were negligible. In addition to the acoustic analysis, several judges listened to the selected audio stimuli to determine the perceived stress and to ensure that all the items sounded natural. These stimuli were then pilot tested with 10 adults, and the items on

TABLE 1
Participant characteristics for Experiment 1

	Williams syndrome (N = 37)		*Learning disabled (N = 36)*		*Normal controls (N = 42)*	
	M (SD)	*Range*	*M (SD)*	*Range*	*M (SD)*	*Range*
Age	19;6 (5.9)	13;0–35	18;0 (2;5)	13;7–23;1	19;7 (5;6)	12;8–32;3
Full Scale IQ	69.4 (12.2)	49–94	68.8 (10.8)	52–92	106.1 (14.8)	86–141*
Verbal IQ	75.7 (10.9)	58–101	76.2 (8.6)	55–97	104 (13.1)	81–132*
Nonverbal IQ	69 (13.9)	40–94	67.1 (15.2)	41–94	107.4 (15.1)	84–140*
Vocabulary	80.2 (9.5)	61–103	82.2 (11.1)	54–102	103.7 (17.1)	82–141**

* $N = 16$ ** $N = 22$

which the pilot participants consistently made errors were dropped from the experiment. The final stimulus set included 11 test word-pairs with double meaning and 12 foils. The ambiguous equal stress recordings and the foils were included to mask the experimental manipulation, and to ensure participants understood the task requirements. Foils were pronounced in two conditions: with their corresponding natural stress and with equal stress on both monosyllabic words. The complete set of stimulus word-pairs is presented in Table 2.

Each test word-pair was presented twice in each of the three conditions: with a high pitch accent on the first syllable (e.g., HOT dog), on the second syllable (hot DOG), or equal stress throughout (HOT DOG), and each foil appeared at least twice, for a total of 95 pseudo-randomised trials. The same word-pair never appeared on successive trials. Two pictures displayed simultaneously on the right and the left side of the computer screen appeared 750 ms before the onset of each audio stimulus. The pictures remained on the screen during the presentation of the sound file and until the participant made a response by pressing one of two buttons (right side, left side) on a button box. For the test word-pairs each picture illustrated one of the two possible lexical interpretations; Figure 1 presents examples of the pictures.

For the foils, one of the pictures was an appropriate illustration of the word-pair, while the second illustrated items somehow related, but with a different meaning (e.g., for *bookcase* there was a picture of a bookshelf with books on it and a picture of a backpack full of books). All pictures were drawn by an amateur cartoonist and coloured with Adobe Photoshop. To ensure that the pictures were clear and unambiguous illustrations of the two meanings of the target items (word-pairs), ten raters were asked to provide

TABLE 2
List of stimulus word-pairs for Experiment 1

Test stimuli	*Foils*
Black Board	Ash Tray
Bulls Eye	Soft Drink
Green House	Blue Bird
High Light	Tee Shirt
Hold Up	Top Shelf
Hot Dog	Head Phones
Make Up	Mail Box
Pick Up	Door Mat
Take Off	Tree House
Top Hat	Book Case
Wet Suit	Wood Pile
	Drive Way

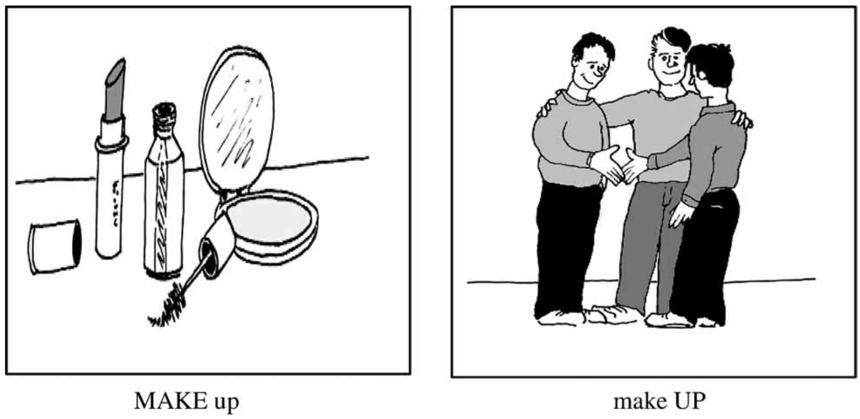

Figure 1. Examples of picture stimuli for Experiment 1.

labels or very brief descriptions of the images, without having heard any audio stimuli, and to indicate how sure they were of their interpretation (i.e., not sure, somewhat sure, very sure). When the images were displayed in the respective pairs corresponding to the target two-syllable constructs, raters provided the matching compound or two-word phrase used in the task on all items, and indicated high certainty about their interpretation. Two images (corresponding to 'wet suit' and 'hold up'), which did not receive maximum certainty ratings, were modified to better illustrate the intended meanings of the lexical constructs. Two orders of stimulus presentation were defined and counterbalanced across participants.

Procedure. The task was programmed using Superlab software and administered on a Dell Latitude laptop computer. Before administering the experiment, training was provided to ensure that the participants understood the task. In training, a new set of word-pairs with double meaning was first introduced using sentences, to facilitate the correct picture selection (e.g., 'Take a copy of the HANDout' and 'If you put your hand OUT, the dog will shake it'). After the sentence presentation, the same word-pair was played again without the sentential context while the same pair of pictures was displayed on the screen. The pictures were switched with respect to their left-right position on the screen and corrective feedback was given as needed. The training consisted of 2 pairs of sentences and 6 word-pairs that were presented in isolation, including word-pairs with double meaning as well as foils. Thus, the training was used to illustrate how prosody can change the meaning of words, and to help the participants become aware of the ambiguities in the stimuli and of how stress placement can change meaning.

For the experiment, participants were told that they would hear some words played by the computer and would see two pictures on the computer screen, side by side, at the same time as the words were played. They were instructed to listen carefully to the way the speaker said the words, and to choose from the two pictures displayed simultaneously on the screen the one that best represented the meaning of the words they heard. The forced choice response was given by pushing one of two buttons on a Cedrus 410 button box, corresponding to the two response options.

Results

Correct performance on the test stimuli presented with either first or second syllable stress is shown in Figure 2. (Because of the inherent ambiguity of the equal stress condition, these stimuli were not included in the analysis of correct performance data.) We conducted a two-way ANOVA with group as between-subjects factor and stress-pattern as within-subjects variable. The main effects for stress-pattern, $F(1, 112) = 84.2$, $p < .001$, partial $\eta^2 = .43$ and group, $F(2, 112) = 64.6$, $p < .001$, partial $\eta^2 = .53$ were significant, as well as the interaction between stress-pattern and group, $F(2, 112) = 9.63$, $p < .001$, partial $\eta^2 = .15$. Post hoc comparisons (Scheffé tests) indicated that the NC group was significantly more accurate than the WS and LID groups in using stress-pattern cues to select the intended meaning ($p < .001$ for both comparisons). The WS and LID groups did not differ from each other in overall accuracy ($p = .82$). The interaction was further analysed by examining performance as a function of stress-pattern within each group separately. In each group the stress pattern effect on accuracy was significant, $F(1, 41) =$

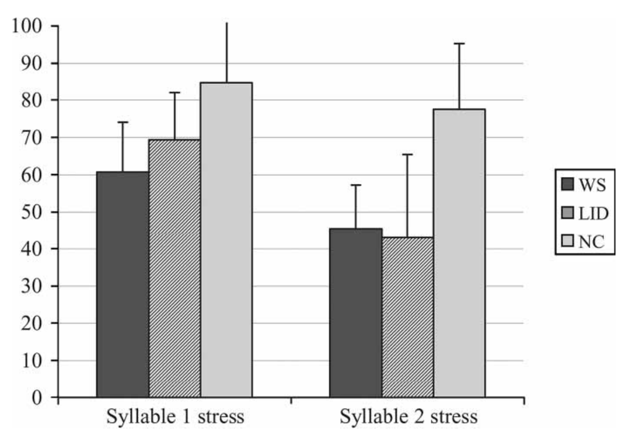

Figure 2. Accuracy in lexical ambiguity resolution as a function of word-pair stress-pattern – Experiment 1.

9.99, $p < .003$, partial $\eta^2 = .19$ for the NC, F (1, 36) = 27.79, $p < .001$, partial $\eta^2 = .44$ for the WS, and $F(1, 35) = 43.62$, $p < .001$, partial $\eta^2 = .56$, for the LID, resulting from better performance on the 1st syllable stress than on the 2nd syllable stress pattern (see Figure 2).

However, for the NC group the difference in mean performance on 1st syllable and 2nd syllable stress items was smaller than the respective difference found in the WS and the LID groups, as indicated by the different magnitudes of the effect size in each group. The WS and LID groups showed a stronger tendency to revert to the 1st syllable stress interpretation (the prevalent pattern in English) when listening to word-pairs spoken with stress on the 2nd syllable, which suggests a lack of clear reliance on stress cues to word meaning.

We also conducted an item analysis ranking each word-pair in order of accuracy within each group, and comparing these rankings across groups. When examining rankings for all 22 items taking into account their stress pattern, only one item had the same rank in all three groups and five word-pairs had the same rank in the WS and the LID groups. Spearman rank-order correlations for item rankings between the groups were mixed: rankings for the WS and LID groups were correlated, $r_s = .58$, $p < .01$, as were those for the NC and LID groups, $r_s = .54$, $p < .01$, while for the WS and the NC groups correlations were not significant, $r_s = .38$, $p = .09$, suggesting that patterns of performance were not entirely item-bound, but reflected specific difficulties in the use of lexical prosodic cues to meaning interpretation in the WS and LID groups.

Accuracy was high on foils in all groups (88–95%), indicating that the majority of participants were able to attend to the task requirements. However, to control for possible performance limitations due to attentional difficulties, we also conducted the same statistical analyses after screening out participants with three or more errors (more than 10% of trials) on foils. This selected group included 22 WS, 30 LID, and 35 NC participants. Results were similar to those obtained in analyses with all the participants from the three groups.

In light of the evidence suggesting that the ability to use stress information to disambiguate word meaning emerges after age 10 (Vogel & Raimy, 2002), we conducted the same analyses on a subgroup of participants who had verbal mental ages of over 12 years based on their PPVT-III scores. This subgroup included 8 participants with WS, 11 with LID, and 32 NC. Accuracy and RT patterns did not differ in any significant way from those found in the larger sample, suggesting that the difficulties encountered by the WS and LID participants in using prosodic cues for lexical decisions may not simply reflect a developmental delay in an otherwise typical process of developing receptive prosody skills.

Discussion

As predicted, the NC participants used stress information consistently in deciding on the matching picture, but were more accurate on the 1st syllable stressed word-pairs than on the 2nd syllable stressed word-pairs. This response pattern suggests that the frequency of typical stress-patterns in a language influences the interpretation of lexically ambiguous word-pairs. The WS and LID groups were less accurate than the NC group in every condition; however, these groups did not differ from each other on accuracy in this experiment. The difference in accuracy between the 1st and 2nd syllable stress word-pairs was greater in both the WS and the LID groups compared with the typical controls, suggesting a tendency in these groups to 'default' to the 1st syllable stress pattern interpretation, the prevalent stress pattern in English, while ignoring the prosodic cues to lexical meaning. Although our results cannot rule out a developmental delay explanation for the difficulties of the WS and LID participants on this task, we found no age-related differences in performance in either group with intellectual disabilities.

EXPERIMENT 2: EVALUATION OF EMOTIONAL PROSODY AND EMOTIONAL CONTENT

This experiment was designed to test explicit recognition of the emotional valence (happy, sad, neutral) of either vocal intonation or sentence meaning in two tasks, administered in counterbalanced order.

Method

Participants. Participants included adolescents and adults with WS ($N = 47$), LID ($N = 37$) and NC ($N = 47$). Many of the same participants completed both Experiments 1 and 2. As in Experiment 1, all three groups were matched on age, $F(2, 128) = 0.74$, $p = .48$, and the WS and LID groups were also well matched on IQ ($p = .98$) and verbal knowledge scores ($p = .62$) on Scheffé post-hoc comparisons.

Stimulus materials. Audio stimuli consisted of 18 target sentences, half with happy content and half with sad content (e.g., *If Mike wins the game, all his friends cheer; When Jane lost her puppy, she started to cry*). Based on the literature regarding acoustic properties of emotion in speech (Bachorowski & Owren, 1995; Banse & Sherer, 1996; Cosmides, 1983; Murray & Arnott, 1993), each target sentence was recorded onto a laptop computer by a speaker trained in voice and prosodic theory in each of three conditions: happy prosody (high pitch, fast rate, ends in complex tone), sad prosody (lower pitch, slower, ends in low tone), and neutral (mid-range pitch, accent on main topic, ends in less complex final tone). Congruent (e.g., happy sentence content + happy tone of voice) and incongruent (e.g., happy sentence content + sad or neutral tone of voice) combinations of semantic content and affective prosody were created and presented in pseudo-randomised order, so that the same sentence did not appear in succession. All sentences were similar in length and syntactic form, and acoustic analyses were conducted (using PRAAT) to validate the prosodic manipulations. In addition, the stimuli were pilot tested on 10 raters to confirm that the vocal intonation conveyed the intended emotions.

For each of the two tasks (Content judgement and Prosody judgement) three versions with different orders of item presentation were created, consisting of 18 sentences each, and these versions were counterbalanced across participants.

Procedure. In the Content judgement task participants were instructed to decide whether the sentence heard was about something happy, or sad. The examiner emphasised to participants that they should listen to the words and ignore the tone of voice of the speaker. In the Prosody judgement task participants were instructed to decide whether the speaker sounded happy, sad, or neither. They responded by pressing one of three clearly marked buttons on a button box. It was emphasised that they should listen just to the speaker's voice and to ignore the words she was saying. A training activity preceded each task to ensure that the participants understood the instructions and the task requirements. Six new utterances used for training were recorded by a different speaker and pronounced with exaggerated emotional prosody to draw attention to the distinction between tone of voice and sentence content. Feedback was provided on these stimuli during training.

The experiment was run using Superlab software and was administered individually in a quiet room. The order of presentation of the Content and Prosody judgement tasks was counterbalanced across participants. Accuracy data were collected using a Cedrus-450 button box and Superlab software.

Results

The performance of the three groups on the Content judgement and the Prosody judgement tasks, for congruent and incongruent combinations of prosody and content is shown in Figure 3.

In judging the emotional content of the sentences all groups performed well, with over 95% accuracy for the trials with matching content and prosody. A two-way ANOVA with group as between-subjects factor and type of prosody/content combination (congruent vs. incongruent) as repeated measure yielded significant main effects of group, $F(2, 128) = 4.27$, $p < .02$, partial $\eta^2 = .06$ and of type of prosody/content combination,

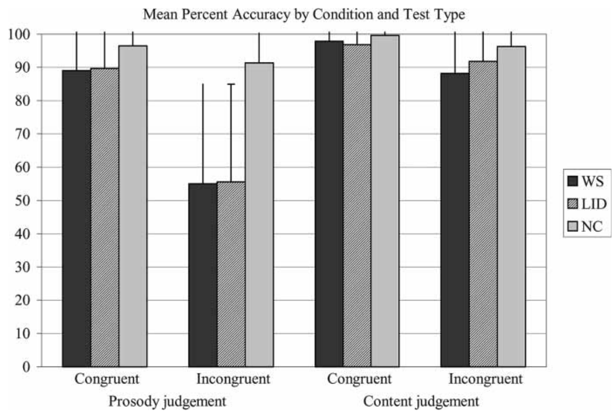

Figure 3. Explicit judgements of affective prosody and affective sentence content – Experiment 2.

$F(1, 128) = 19.2$, $p < .001$, partial $\eta^2 = .13$, with no significant interaction. Post hoc Scheffé tests indicated that the NC group was overall significantly more accurate than the WS group in judging emotional content ($p < .03$). No other group differences were significant. All participants were more accurate in the congruent compared to the incongruent condition when judging the Content of the sentences. However, this difference was significant for the NC group, $F(1, 46) = 5.58$, $p < .03$, partial $\eta^2 = .11$ and for the WS group, $F(1, 46) = 14.42$, $p < .001$, partial $\eta^2 = .23$, but only marginally significant for the LID group, $F(1, 36) = 2.88$, $p = .098$, partial $\eta^2 = .07$.

For the Prosody judgement task, a similar two-way ANOVA yielded significant main effects of group, $F(2, 128) = 35.8$, $p < .001$, partial $\eta^2 = .35$; of type of combination prosody/content, $F(1, 128) = 115.7$, $p < .001$, partial $\eta^2 = .47$; and a significant interaction between group and combination type, $F(2, 128) = 18.9$, $p < .001$, partial $\eta^2 = .22$. The NC group was significantly more accurate than either of the clinical groups in judging emotional prosody ($p < .001$, both comparisons, Scheffé tests), whereas the WS and LID groups did not differ from each other ($p = .99$). The significant interaction was further analysed by assessing group differences separately for each type of prosody/content combination. In the congruent condition, although there was a significant group effect, $F(2, 128) = 5.87$, $p < .01$, the effect size, partial $\eta^2 = .08$, was small. Post hoc Scheffé tests indicated that the WS and LID groups performed similarly in the congruent condition ($p = .97$), whereas the NC group was more accurate than both the WS group ($p < .01$) and the LID group ($p < .05$). In the incongruent condition there was a significant group effect, $F(2, 128) = 32.9$, $p < .001$, and a larger effect size, partial $\eta^2 = .34$. Post hoc Scheffé tests indicated that the WS and LID groups performed significantly worse on the incongruent stimuli than the NC group ($p < .001$), but that these two groups did not differ from each other ($p = .96$).

We further examined the magnitude of the interference effect of semantic information on judgements of affective prosody within each group separately. As predicted, in all groups accuracy was significantly higher in the congruent condition compared to the incongruent condition (see Figure 3), but the relative effect size of the prosody/content combination type was larger for the WS group, $F(1, 46) = 57.9$, $p < .0001$, partial $\eta^2 = .56$, and the LID group, $F(1, 36) = 42.4$, $p < .0001$, partial $\eta^2 = .54$, compared to the NC group, $F(1, 46) = 10.4$, $p < .01$, partial $\eta^2 = .18$.

Thus, all groups showed an interference effect from the channel-to-be-ignored on their explicit judgements of either vocal prosody or semantic content, as suggested by the lower accuracy in the incongruent condition on both tasks. But which type of information (prosodic or semantic) has a stronger interference effect, and is the interference pattern the same in the three groups? We addressed these questions by comparing performance on the two tasks (the two types of explicit judgements) for the incongruent sentence combinations. A two-way ANOVA with group as between-subjects and type of explicit judgement (prosody vs. content) as within-subjects variable yielded a significant effect of type of judgement, $F(1, 128) = 93.1$, $p < .001$, partial $\eta^2 = .42$, and a main effect of group, $F(2, 128) = 33.1$, $p < .001$, partial $\eta^2 = .34$, as well as a significant interaction between type of judgement and group, $F(2, 128) = 15.5$, $p < .001$, partial $\eta^2 = .19$. All three groups performed better in judging the emotional content than the emotional prosody information in incongruent sentences, $F(1, 46) = 6.77$, $p < .02$, partial $\eta^2 = .13$ for the NC group, $F(1, 46) = 43.3$, $p < .0001$, partial $\eta^2 = .49$ for the WS group, and $F(1, 36) = 37.6$, $p < .0001$, partial $\eta^2 = .51$ for the LID group, consistent with the expectation of a content-bias in the interpretation of speech. However, the interference effect of content information on judgements of vocal prosody was larger than the interference of vocal prosody information on judgements of semantic content for the WS and the LID groups, as indicated by the group differences in effect sizes (see Figure 3).

Thus, both the WS and the LID participants had significant difficulties ignoring the content information when having to judge the emotional tone of voice, while the NC group was able to judge the tone of voice in the presence of conflicting content information. Nevertheless, explicit content judgements appeared to be easier than explicit vocal prosody judgements in the presence of an incongruent channel of information even for the NC group, although the difference in accuracy for the two types of explicit judgements was significantly smaller for the typical controls than for the clinical groups.

Discussion

As predicted, the NC group performed more accurately on both the affective Prosody and the affective Content judgement tasks. On both tasks accuracy was lower in the incongruent conditions suggesting that the presence of a conflicting channel information interfered with the participants' ability to focus on one or the other of the information sources. This effect was significantly stronger for the prosody judgement task, particularly for the WS and LID participants, suggesting a strong content bias in listening to speech. Contrary to our predictions, the participants with WS did not show special sensitivity to affective prosody in this experiment.

EXPERIMENT 3: EVALUATION OF EMOTIONAL PROSODY IN FILTERED SPEECH

It is possible that tasks requiring judgements of vocal prosody in the presence of conflicting content, as used in Experiment 2, entail significant attentional task demands that may have masked greater sensitivity to affective prosody in the WS group. To explore this possibility we administered a task that required participants to identify the emotion expressed in the tone of voice of the speaker when the speech content was not comprehensible, by using low-pass filtered speech.

Method

Participants. A subgroup of the individuals who had participated in Experiments 1 and 2 (36 participants with WS, 36 with LID and 35 NC) participated in Experiment 3 several months after having completed Experiment 2. As before, the three groups were well matched on age, $F(2, 104) = .23$, $p = .79$, and the WS and LID groups were also matched on IQ ($p = .99$) and PPVT scores ($p = .71$).

Stimulus materials. The audio stimuli were the same sentences used in Experiment 2, which were low-pass filtered, removing all the acoustic information above 100–150 Hz (higher frequencies being related to phonemes and lexical content), but retaining the acoustic information that determined the affective prosody of the utterances. Thus, the content information of the filtered speech was incomprehensible, while the acoustic cues of emotional prosody were preserved. All stimuli were subjected to acoustic analysis of pitch and duration that validated the prosodic manipulations, and were also pilot tested by naive listeners to be certain that no words could be understood. Three versions of the task, each consisting of 18 sentences arranged in three different pseudo-randomised orders, were created and administered, counterbalancing versions across participants.

Procedure. A training activity with feedback preceded the administration of the experiment. Six utterances recorded by a different speaker with exaggerated emotional prosody were low-pass filtered and used for training. Feedback was provided on the six sample stimuli during training, pointing out how the tone of voice could convey the speaker's emotions, even when the content of speech was incomprehensible. Participants were told that they would not be able to understand the words played by the computer, and were instructed to identify the emotion of the speaker by pressing one of the three clearly marked buttons on a button box, corresponding to Happy, Sad, Neither.

Results

Correct performance on each emotion for each group is presented in Figure 4. A group by emotion ANOVA on accuracy data yielded a main effect of group, $F(2, 104) = 14.16$, $p < .001$, $\eta^2 = .21$; of emotion, $F(2, 103) = 53.5$, $p < .001$, $\eta^2 = .51$, but no significant interaction. Pairwise comparisons (Scheffé tests) indicated that the WS group was significantly more accurate than the LID group ($p < .01$), but not significantly different from the NC group ($p = .20$) in identifying affective prosody in filtered speech. In all groups accuracy was significantly better for happy and sad vocal prosody compared to neutral tone of voice ($p < .001$, both comparisons), and there were no differences in accuracy between happy and sad utterances ($p = .95$).

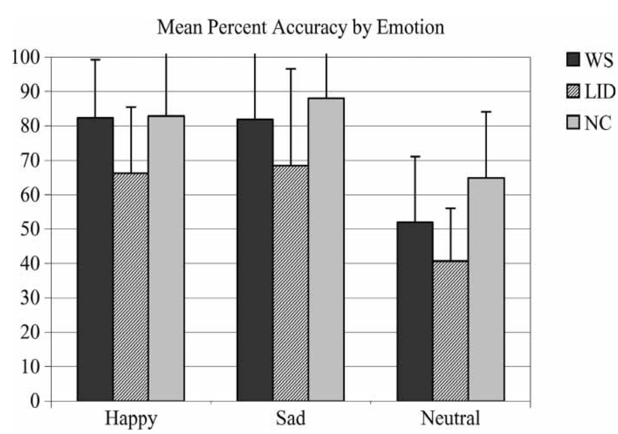

Figure 4. Judging affective prosody in filtered speech – Experiment 3.

Discussion

The task used in this experiment is the same as the Prosody judgement task used in Experiment 2, but with the essential modification that this time prosody was not pitted against semantic content. Thus, this experiment provides a clearer measure of sensitivity to affective prosody in the absence of competing task demands in language. In this context, the WS participants performed close to the accuracy level of the age-matched NC group, and significantly better than the IQ and vocabulary-matched LID group. Performance differences in the two clinical groups on the filtered speech task suggest relative sparing in identifying affective prosody in WS in the absence of competing linguistic demands.

GENERAL DISCUSSION

The main goal of the studies reported in this paper was to investigate receptive prosody in adolescents and adults with WS, comparing this group with age-matched controls and with a second group of age, IQ and language matched participants with LID. Previous research has demonstrated that people with WS are significantly more expressive in narrative and conversational discourse than IQ-matched groups with Down syndrome and mental age matched controls (Reilly et al., 1990, 2004), although their language skills tend to be commensurate with mental age (Mervis, Morris, Bertrand, & Robinson, 1999). We hypothesised that the participants with WS would perform at the same relatively impaired level as the LID group on linguistic prosody judgements but would show spared ability to interpret affective prosody by performing significantly better than the LID group and at the same level as typical controls. These hypotheses were generally supported by the findings from Experiments 1 and 3 reported in this paper; however, the contrasting findings from Experiments 2 and 3 demonstrated that people with WS only show sensitivity to affective prosody when no linguistic information is included in the auditory stimulus. The relative sparing of affective prosody coupled with relative impairment in linguistic prosody found in the WS group suggest that these aspects of prosody may be dissociated in this population.

In the first experiment, both clinical groups had difficulties using stress patterns to determine the intended meaning when listening to dual-meaning word-pairs. They were less sensitive to linguistic prosodic information than the NC group, as evidenced by their lower accuracy in selecting the intended meaning of two-syllable lexical constructs based on stress pattern information. At the same time, all three groups showed a bias toward better performance on word-pairs with first syllable stress, which is the dominant stress pattern for English (Cutler & Carter, 1987). These findings indicate that all the groups were sensitive to stress patterns in their native language. Studies on typically developing children suggest that mastery of linguistic prosody is not fully achieved until early adolescence (Vogel & Raimy, 2002). The relatively poor performance by the participants with WS and LID in Experiment 1 may suggest arrested development, related to their more limited language ability. Although our analysis of data from the subgroup of participants with verbal age levels above 12 years did not reveal more adult-like performance, this might be because the number of participants in these subgroups was quite small, or because verbal age was estimated from the participants' scores on the PPVT-III, which generally overestimates overall linguistic ability, especially in WS (Mervis et al., 1999). Another explanation for the poorer performance by the clinical groups could be their difficulty integrating attention to the acoustic signal while selecting between two picture choices. On this view, the linguistic prosody task we used had challenging attentional demands that may have affected overall performance. Support for this interpretation comes from the overall performance by the NC group that was below ceiling level. Further studies are needed to distinguish between these interpretations of our findings.

Our findings from the second experiment did not support our initial hypothesis that the WS group might be especially sensitive to prosodic cues about a speaker's emotional state. For the Prosody judgement task, participants had to make explicit judgements of emotional prosody when vocal prosody was combined with either congruent or incongruent semantic content. In the congruent condition, the WS and LID groups performed well, at around 90% accuracy, however in the incongruent condition, both these groups showed significant difficulties compared with the NC group, with performance just above 50% (chance performance would be at 33%). These data suggest a strong content bias in judging a speaker's emotion, even when asked to explicitly focus on the vocal intonation. As in the first experiment, this pattern is similar to performance by children younger than 10 years in prosody experiments, and may partially reflect the high attentional or executive demands of the task (Morton & Trehub, 2001). Difficulties with attention modulation, especially distractibility and concentration problems have been documented in research on the behavioural phenotype of WS using parental or teacher report measures (Dilts, Morris, & Leonard, 1990; Gosh & Pankau, 1994; Greer, Brown, Pai, Choudry, & Klein, 1997; Sarimski, 1997; Udwin & Yule, 1991). Similar problems on executive function tasks that require inhibitory control have been documented in experimental studies with children with WS, such as relatively poor performance on a Stroop-like task developed for use with young children, the Day-Night task (Gerstadt, Hong, & Diamond, 1994). Tager-Flusberg, Sullivan, and Boshart (1997) found that only about 21% of a group of

5–8-year-olds with WS passed the day-night Stroop task, compared with about 50% of a control group with developmental disabilities matched on age and mental age. However, these studies focused on children, so it is not clear to what extent such attention modulation and executive function difficulties persist into adolescence and adulthood. More research is needed to disentangle the possible roles of sensitivity to prosody from attention modulation demands, such as those inherent in the interference-type task in the second experiment.

All the groups in Experiment 2 were influenced by affective prosody cues, when making explicit judgements of the emotional content of sentences, as reflected in lower accuracy in the incongruent condition of the Content judgement task; however, the WS group was not more influenced by incongruent prosody on this task than either of the other groups as we had originally predicted. The bias toward judging content over prosody was found in all the groups, reflected in better performance on the Content judgement task compared with the Prosody judgement task. This pattern reflects the cultural bias for English in which content is the primary means for conveying information (e.g., Kitayama, 2002), and again indicates that all three groups of participants had acquired this cultural-linguistic knowledge of the communication practices associated with their native language.

The difficulty experienced by the WS and LID groups in interpreting the speaker's emotional state from prosodic cues when conflicting semantic content was presented, suggests that these clinical groups would have problems understanding non-literal language, especially sarcasm or irony. Correct interpretation of these types of utterances as used in everyday communication entails the ability to integrate context or contradictory prosodic cues with a literal interpretation of an utterance (Ackerman, 1986). Sullivan, Winner, and Tager-Flusberg (2003) found that fewer than 10% of adolescents with WS or LID were able to discriminate between lies and ironic jokes, even though most of their participants had the requisite theory of mind skills. Although Sullivan and her colleagues did not provide prosodic cues, only contextual cues, these findings highlight the deficits in interpreting ironic jokes in people with developmental disorders, including WS. The findings from Experiment 2 reported in this paper suggest that the clinical groups would have similar difficulty using prosodic information to interpret non-literal language. Problems distinguishing between literal and non-literal language have important consequences for everyday social interactions. The inability to tell if another person is lying or joking would inevitably lead to difficulties in peer interactions and may make a person with WS or LID more vulnerable to teasing or the target of cruel jokes (cf. Davies, Udwin, & Howlin, 1998; Guralnick, 1984; Kasari & Bauminger, 1998).

In contrast to the findings from Experiment 2, the third experiment, which used filtered speech, found that adolescents and adults with WS were relatively spared in their ability to interpret affective prosody when no additional linguistic information was present in the acoustic stimulus. Specifically, on filtered speech the WS participants performed significantly better than the LID group, and almost at the same level as the NC group. These findings suggest that, as we hypothesised, people with WS have special sensitivity to non-linguistic affective information, which is consistent with the view that they are especially empathic toward others (Klein-Tasman & Mervis, 2003; Rowe, Beccera, & Mervis, 2002; Tager-Flusberg & Sullivan, 1999). Nevertheless, it is important to note that we only found this sensitivity to affective prosody when only tone of voice was presented, which is not typical of everyday communication. Moreover, since only two emotions, *happy* and *sad*, were included in Experiment 3 along with *neutral* intonation, perhaps performance of the WS participants was more related to judging positive versus negative valence, than to discerning specific emotions from vocal cues. In the second experiment in which more natural stimuli that included both content and prosody were used, the WS group was no better than the LID group in judging emotional state from prosodic cues. Furthermore, Plesa-Skwerer and her colleagues found that adolescents and adults with WS and LID showed similar impairments relative to typical controls in judging negative emotions (*sad, angry, fearful*) in vocal stimuli on the DANVA-2, a standardised test in which a child or adult uses intonation to express different basic emotions when speaking the same neutral sentence (Plesa-Skwerer, Faja, Schofield, Verbalis, & Tager-Flusberg, 2006). Thus, the findings reported here regarding the special sensitivity to affective prosody in WS have limited implications for how well people with WS are able to use this information in their daily lives.

Overall, our research suggests that people with WS have some difficulty interpreting prosodic information, especially when presented in linguistic contexts. These receptive impairments contrast with the reports in the literature on their spared expressive skills (Harrison et al., 1995; Losh et al., 2000; Reilly et al., 2004). This contradictory profile of receptive and expressive ability suggests that in everyday social interactions, impairments in the ability to interpret social cues may be masked by an overabundance of linguistic and dramatic expressive devices used to engage and involve the social partners or audience and to maintain social contact, especially since people with WS find social interaction highly rewarding (Doyle, Bellugi, Korenberg, & Graham, 2003).

In sum, people with WS present a complex and often contradictory picture of spared and impaired performance on tasks that tap the ability to interpret or express linguistic and affective prosody. To some extent, this picture reflects the complex profile of the social phenotype of this population: children and adults with WS are intensely interested in other people and very socially engaging, but they also have great difficulties in peer relations and making friends (Tager-Flusberg & Plesa-Skwerer, 2005), although we do not know whether performance on prosody tasks is directly related to social adaptation in WS or other populations. Future research should investigate receptive and expressive prosodic skills within the same individuals across a range of contexts and emotions, and explore how prosodic abilities relate to other aspects of nonverbal communication including facial expression or gesture.

Manuscript received November 2005
Revised manuscript received February 2006
First published online 10 May, 2006

REFERENCES

Ackerman, B. P. (1983). Form and function in children's understanding of ironic utterances. *Journal of Experimental Child Psychology*, *35*, 487–508.

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

Bachorowski, J., & Owren, M. J. (1995). Vocal expressions of emotion: Acoustic properties of speech are associated with emotional intensity and content. *Psychological Science*, *6*, 219–225.

Banse, R., & Sherer, K. R. (1996). Acoustic profiles in vocal emotion and expression. *Journal of Personality and Social Psychology*, *79*, 614–636.

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language*, *30*, 663–644.

Bellugi, U., Marks, S., Bihrle, A., & Sabo, H. (1988). Dissociation between language and cognitive functions in Williams syndrome. In D. Bishop & K. Mogford (Eds.), *Language development in exceptional circumstances* (pp. 177–189). London: Churchill Livingstone.

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St. George, M. (2000). The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience*, *12*(Supplement), 7–29.

Bellugi, U., Wang, P., & Jernigan, T. L. (1994). Williams syndrome: An unusual neuropsychological profile. In S. Broman & J. Grafman (Eds.), *Atypical cognitive deficits in developmental disorders: Implications for brain function* (pp. 23–56). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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

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

Cosmides, L. (1983). Invariances in the acoustic expression of emotion during speech. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 864–881.

Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133–142.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech*, *40*, 141–200.

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

Davies, M., Udwin, O., & Howlin, P. (1998). Adults with Williams syndrome. *British Journal of Psychiatry*, *172*, 273–274.

Dews, S., Winner, E., Kaplan, J., Rosenblatt, E., Hunt, M., Lim, K., McGovern, A., Qualter, A., & Smarsh, B. (1996). Children's understanding of the meaning and functions of verbal irony. *Child Development*, *67*, 3071–3085.

Dilts, C., Morris, C. A., & Leonard, C. A. (1990). Hypothesis for a development of a behavioral phenotype in Williams syndrome. *American Journal of Medical Genetics Supplement*, *6*, 126–131.

Doyle, T. F., Bellugi, U., Korenberg, J. R., & Graham, J. (2004). 'Everybody in the world is my friend'. Hypersociability in young children with Williams syndrome. *American Journal of Medical Genetics*, *124A*, 263–273.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test – III*. Circle Pines, MN: American Guidance Services.

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Ferreira, V., & Dell, G. (2000). Effects of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, *40*, 296–340.

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

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3½ – 7 years old on a Stroop-like day-night test. *Cognition*, *52*, 129–153.

Gosh, A., & Pankau, R. (1994). Social-emotional and behavioral adjustment in children with Williams-Beuren syndrome. *American Journal of Medical Genetics*, *52*, 291–296.

Gosh, A., & Pankau, R. (1997). Personality characteristics and behavior problems in individuals of different ages with Williams syndrome. *Developmental Medicine and Child Neurology*, *39*, 527–533.

Greer, M. K., Brown, F. R., Pai, S., Choudry, S. H., & Klein, A. J. (1997). Cognitive, adaptive and behavioral characteristics of Williams syndrome. *American Journal of Medical Genetics*, *74*, 521–525.

Guralnick, M. J. (1984). The peer interaction of young developmentally delayed children in specialized and integrated settings. In T. Field, J. L. Roopnarine, & M. Segal (Eds.), *Friendships in normal and handicapped children* (pp. 139–152). Norwood, NJ: Ablex.

Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A test of Relevance theory. *Cognition*, *48*, 101–119.

Harrison, D., Reilly, J., & Klima, E. (1995). Unusual social behavior in Williams syndrome: Evidence from biographical interviews. *Genetic Counseling*, *6*, 181–183.

Ishii, K., Reyes, J. A., & Kitayama, S. (2003). Spontaneous attention to word content versus emotional tone: Differences among three cultures. *Psychological Science*, *14*, 39–46.

Jones, W., Bellugi, U., Lai, Z., Chiles, M., Reilly, J., Lincoln, A., & Adolphs, R. (2000). Hypersociability in Williams syndrome. *Journal of Cognitive Neuroscience*, *12*(Supplement), 30–46.

Karmiloff-Smith, A., Grant, J., Berthoud, I., Davies, M., Howlin, P., & Udwin, O. (1997). Language and Williams syndrome: How intact is 'intact'? *Child Development*, *68*, 246–262.

Kasari, C., & Bauminger, N. (1998). Social and emotional development in children with mental retardation. In J. A. Burack, R. M. Hodapp, & E. Zigler (Eds.), *Handbook of mental retardation and development* (pp. 411–433). Cambridge, UK: Cambridge University press.

Kaufman, A., & Kaufman, N. (1990). *Manual for the Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service.

Keenan, T. R., & Quigley, K. (1999). Do young children use echoic information in their comprehension of sarcastic speech? A test of echoic mention theory. *British Journal of Developmental Psychology*, *17*, 83–96.

Kitayama, S. (1996). Rememberance of emotional speech: Improvement and impairment of incidental verbal memory by emotional voice. *Journal of Experimental Social Psychology*, *32*, 289–308.

Kitayama, S. (2002). Culture and basic psychological processes: Toward a system view of culture. *Psychological Bulletin*, *128*, 89–96.

Kitayama, S., & Howard, S. (1994). Affective regulation of perception and comprehension. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 41–65). New York: Academic Press.

Kitayama, S., & Ishii, K. (2002). Words and voice: Spontaneous attention to emotional speech in two cultures. *Cognition and Emotion*, *16*, 29–59.

Kjelgaard, M., & Speer, S. (1999). Prosodic facilitation and inhibition in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, *40*, 153–194.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Klein-Tasman, B. P., & Mervis, C. B. (2003). Distinctive personality characteristics of 8-, 9-, and 10-year-olds with Williams syndrome. *Developmental Neuropsychology*, *23*, 269–290.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America*, *32*, 451–454.

Losh, M., Bellugi, U., Reilly, J., & Anderson, D. (2000). Narrative as a social engagement tool: The excessive use of evaluation in narratives from children with Williams syndrome. *Narrative Inquiry*, *10*, 265–290.

McCann, J., & Peppe, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language and Communication Disorders*, *38*, 325–350.

Mervis, C. B., Morris, C. A., Bertrand, J., & Robinson, B. (1999). Williams syndrome: Findings from an integrated program of research. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 65–110). Cambridge, MA: MIT Press.

Mervis, C. B., Robinson, B. F., Bertrand, J., Morris, C. A., Klein-Tasman, B. P., & Armstrong, S. C. (2000). The Williams syndrome cognitive profile. *Brain and Cognition*, *44*, 604–628.

Morris, C. A., & Mervis, C. B. (1999). Williams syndrome. In S. Goldstein & C. R. Reynolds (Eds.), *Handbook of neurodevelopmental and genetic disorders in children* (pp. 555–590). New York: Guilford.

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

Morton, B. J., Trehub., S. E., & Zelazo, P. D. (2003). Sources of inflexibility in 6-year-olds' understanding of emotion in speech. *Child Development*, *74*, 1857–1868.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*, *93*, 1097–1108.

Nespor, I., & Vogel, J. (1987). *Prosodic phonology.* Dordrecht: Foris.

Nygaard, L., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory and Cognition*, *30*, 583–593.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *35*, 205–220.

Plesa Skwerer, D., Faja, S., Schofield, C., Verbalis, A., & Tager-Flusberg, H. (2006). Perceiving facial and vocal expressions of emotion in Williams syndrome. *American Journal on Mental Retardation*, *111*(1), 15–26.

Reilly, J., Klima, E., & Bellugi, U. (1990). Once more with feeling: Affect and language in atypical populations. *Development and Psychopathology*, *2*, 367–391.

Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). 'Frog, where are you?' Narratives in children with specific language impairment, early focal brain injury and Williams syndrome. *Brain and Language*, *88*, 229–247.

Rowe, M., Beccera, A., & Mervis, C. B. (2002). *The development of empathy in 4-year-old children with Williams syndrome*. Paper presented at the 9th International Professional Conference on Williams Syndrome. Long Beach, CA.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: a study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders*, *32*, 189–194.

Sarimski, K. (1997). Behavioral phenotypes and family stress in three mental retardation syndromes. *European Child and Adolescent Psychiatry*, *63*, 26–31.

Schirmer, A., Kotz, A. S., & Friederici, A. D. (2003). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research*, *14*, 228–233.

Sullivan, K., Winner, E., & Tager-Flusberg, H. (2003). Can adolescents with Williams syndrome tell the difference between lies and jokes? *Developmental Neuropsychology*, *23*, 85–103.

Tager-Flusberg, H., & Plesa-Skwerer, D. (2005). Social engagement in Williams syndrome. In P. J. Marshall & N. A. Fox (Eds.), *The development of social engagement: Neurobiological perspectives* (pp. 331–354). New York: Oxford University Press.

Tager-Flusberg, H., & Sullivan, K. (1999, April). *Are children with WS spared in theory of mind?* Paper presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, NM.

Tager-Flusberg, H., Sullivan, K., & Boshart, J. (1997). Executive functions and performance on false belief tasks. *Developmental Neuropsychology*, *13*, 487–493.

Udwin, O., & Yule, W. (1991). A cognitive and behavioral phenotype in Williams syndrome. *Journal of Clinical and Experimental Neuropsychology*, *13*, 232–244.

Vogel, I., & Raimy, E. (2002). The acquisition of compound vs. phrasal stress: The role of prosodic constituents. *Journal of Child Language*, *29*, 225–250.

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