Reduced Auditory Processing Capacity during Vocalization in Children with Selective Mutism

Miri Arie, Yael Henkin, Dominique Lamy, Simona Tetin-Schneider, Alan Apter, Avi Sadeh, and Yair Bar-Haim

Background: *Because abnormal Auditory Efferent Activity (AEA) is associated with auditory distortions during vocalization, we tested whether auditory processing is impaired during vocalization in children with Selective Mutism (SM).*

Methods: *Participants were children with SM and abnormal AEA, children with SM and normal AEA, and normally speaking controls, who had to detect aurally presented target words embedded within word lists under two conditions: silence (single task), and while vocalizing (dual task). To ascertain specificity of auditory-vocal deficit, effects of concurrent vocalizing were also examined during a visual task.*

Results: *Children with SM and abnormal AEA showed impaired auditory processing during vocalization relative to children with SM and normal AEA, and relative to control children. This impairment is specific to the auditory modality and does not reflect difficulties in dual task per se.*

Conclusions: *The data extends previous findings suggesting that deficient auditory processing is involved in speech selectivity in SM.*

Key Words: Auditory Processing, elective mutism, selective mutism, social anxiety, social phobia, vocalization

Selective Mutism (SM) is a psychiatric disorder of childhood characterized by consistent failure to speak in specific social situations (where speaking is expected) despite ability to speak normally in other situations (DSM-IV-TR). Based on the literature describing the connections between vocalization and hearing, and on review of the neural substrates supporting these connections, Bar-Haim et al (2004) suggested that deficient auditory processing during vocalization might impair the ability of some children with SM to simultaneously speak and process incoming auditory signals.

To preserve auditory sensitivity to external signals during stimulation by one's own voice, negative feedback circuits are activated (Hoy 2002). In humans, activity of the middle-ear acoustic reflex (MEAR) during vocalization results in decreasing the masking influence of the speaker's own voice, thereby improving the speaking person's ability to hear external sounds while vocalizing (Borg et al 1984; Borg and Zakrisson 1973; Borg and Zakrisson 1975). Additionally, efferent activity of the medial olivocochlear bundle has been implicated in the improvement of signal-to-noise ratio and speech intelligibility in noise (Dewson 1968; Giraud et al 1997; Micheyl and Collet 1996), and thus has an important role in preventing desensitization of the auditory system during vocalization. Bar-Haim et al (2004) reported that, compared with speaking control children, children with SM displayed significant aberrations in MEAR thresholds and decay functions, and diminished suppression effect of transient oto-acoustic emissions (TEOAE), indexing reduced activity of efferents from the olivocochlear bundle.

The objective of the present study was to test whether auditory processing is indeed impaired during vocalization in children with SM who display abnormal auditory efferent activity (AEA), with a new sample of children. Specifically, while all children were expected to show poorer performance in a task requiring both auditory processing and vocalizing (dual task) relative to a task requiring auditory processing alone (single task), we expected that relative to children with SM who have normal AEA, and relative to normally speaking controls, children with SM and abnormal AEA would show a greater dual-task performance cost. Furthermore, we expected these results to be specific to auditory-vocal performance, and not to occur on a visual-vocal task.

Methods and Materials

Participants

Participants were 28 children recruited into three study groups: 9 children with SM and abnormal AEA; 9 children with SM and normal AEA; and 10 speaking control children. Table 1 summarizes children's characterization data by group. For complete description of the referral process, diagnostic procedures, inclusion criteria, and questionnaires used in the study, see Supplement 1.

Audiologic Assessment

Children who met the study's psychiatric inclusion criteria were invited for audiologic assessments. Children with normal air-conduction thresholds (i.e., pure-tone average of 0.5, 1, and 2 kHz $\leq$ 15dBHL), normal tympanograms, and normal auditory brainstem response (ABR) were further tested for AEA function.

The assessment of AEA function included testing of ipsi- and contra-lateral MEAR pure-tone thresholds at 0.5, 1, and 2 kHz, ipsi- and contra-lateral reflex decay to 0.5 and 1 kHz pure tones, and TEOAE suppression effect in both ears. For detailed description of audiologic procedures and criteria for auditory efferent deficiency, see Bar-Haim et al (2004) and Supplement 1.

Assessment of Auditory Monitoring Performance During Vocalization

All children were trained to vocalize (counting 1 to 10 repeatedly) until preset criteria of fluency (breaks of silence not

From the Adler Center for Research in Child Development and Psychopathology, Department of Psychology (MA, DL, AS, YB), Tel Aviv University; Department of Communication Disorders (YH, STC), Sackler Faculty of Medicine, Tel Aviv University; the Speech and Hearing Center (YH, STC), The Chaim Sheba Medical Center; Department of Psychiatry (AA), Sackler Faculty of Medicine, Tel Aviv University; the Feinberg Child Study Center (AA), Schneider's Children's medical Center of Israel.

Address reprint requests to Yair Bar-Haim, Ph.D., The Adler Center for Research in Child Development and Psychopathology, Department of Psychology, Tel-Aviv University, Ramat Aviv, Tel-Aviv 69978, Israel; E-mail: yair1@post.tau.ac.il.

Received August 10, 2005; revised February 8, 2006; accepted February 10, 2006.

0006-3223/07/$32.00
doi:10.1016/j.biopsych.2006.02.020

© 2007 Society of Biological Psychiatry

Table 1. Characteristics of Study Participants

	Selective Mutism Abnormal Efferents	Selective Mutism Normal Efferents	Normal Controls
Age (years)	9.06 (2.04)	8.72 (1.91)	9.31 (1.21)
Sex (No. male/female)	3/6	5/4	7/3
Co-morbidity, (No.)			
Social Phobia	5	3	—
Separation Anxiety	1	—	—
Dysthemia	—	1	—
ADHD	1	1	—
Enuresis	2	1	—
SMQ	17.56 (7.68) a	17.44 (7.58) a	39.00 (7.58) b
SPAI-C	24.19 (7.89) a	18.67 (10.98) a	7.41 (7.11) b
SCARED-C	28.67 (9.57) a	20.44 (8.50)	15.70 (7.23) b
SCARED-P	21.67 (7.05) a	20.11 (11.34) a	7.30 (5.27) b
CBCL – Internalizing	.39 (.87) a	.21 (1.21) a	−1.04 (.56) b
CBCL – Externalizing	−.75 (.47)	−1.16 (.43)	−1.20 (.55)

Means and standard deviations in parentheses unless otherwise specified.
Different lower case letters represent significant post-hoc Bonferroni contrasts, $p < .05$.
SMQ = Selective Mutism Questionnaire; SPAI-C = Social Phobia and Anxiety Inventory for Children; SCARED-C/P = Screen for Child Anxiety Related Emotional Disorders – Child/Parent report; CBCL = Child Behavior Checklist (0 = standardized mean score of a clinically referred population).

exceeding 2 sec), loudness (>30 dB SPL), and duration (10 sec) were reached. Children received visual feedback; a green light when vocalizing within criterion parameters and a red light when failing to do so. There was considerable variability in the time it took children to reach the vocalizing criterion. However, only one child failed to reach criterion and was excluded from the study.

The experiment included two single-task conditions: Auditory monitoring (detecting a target word in a list of spoken words) and visual monitoring (detecting a target picture in a series of pictures), and two corresponding dual-task conditions, in which a vocalization task (repeatedly counting 1 to 10) was added to the single-task conditions. In the dual-task conditions, vocalization was the primary task.

In the auditory tasks, children had to press a key upon detection of a target word randomly embedded between the 3rd and 14th locations in a list of 16 words presented at a rate of one per sec. Each child listened to ten different lists of words delivered binaurally at 65 dB HL. One of these lists contained no target word and served as a catch trial. Trial presentation order was randomized. The words in each trial were randomly selected from a pool of 60 phonetically balanced monosyllabic words. These words were recorded by a male speaker in a soundproof room, with sampling rate of 22,050 Hz, and 16 bits quantization level. The words were of similar duration and their amplitudes normalized.

In the visual tasks, children observed 10 sets of pictures consisting of 16 monochromatic images each. The pictures were 12 × 12-cm visual representations of the words used in the auditory tasks, with random noise pixilation added to them. Children had to press a key upon detection of a target picture appearing between the 3rd and 14th locations in the trial. One catch trial was included. Pictures were presented at the center of a 17″ computer screen for 150 ms and then masked with a checkerboard image that remained on the screen for 1850 ms, until the next picture was displayed.

Single tasks were presented first and were followed by dual tasks. Differences in practice between the single- and dual-task conditions were not a concern because dual-task performance was expected to yield poorer performance despite increased practice. Order of task presentation by modality was counterbalanced within each condition (single/dual) and group. Performance accuracy was measured as percent errors in each condition.

Results

Participants' error rates on the monitoring tasks were analyzed via repeated measures ANOVA with Modality (auditory, visual) and Type of Task (single, dual) as within-subject factors, and Group (SM abnormal AEA, SM normal AEA, normal control) as a between-subjects factor. Results are summarized in Figure 1. A main effect of Type of Task, $F(1, 25) = 33.20$, $p < .0001$ was qualified by a Modality by Type of Task interaction, $F(1, 25) = 17.63$, $p < .0001$, and a nearly significant Modality by Type of Task by Group three-way interaction, $F(2, 25) = 2.61$, $p = .08$. No other effects approached significance.

Separate follow-up ANOVAs for each Modality revealed a Type of Task by Group interaction for the auditory tasks, $F(2, 25) = 4.14$, $p < .05$, and no such interaction for the visual tasks, $F(2, 25) = .23$, $p = .79$. Post-hoc contrasts showed that children from the three groups did not differ in performance on the single auditory task. However, when required to vocalize in the dual auditory task, children with SM and abnormal AEA committed more errors than children with SM and normal AEA, $t(16) = 2.53$, or children in the control group, $t(17) = 2.13$, $ps < .05$.

Discussion

The results show that the ability of children with SM and abnormal AEA to process auditory input is impaired during vocalization relative to children with SM who have normal AEA, and relative to speaking control children. These findings are specific to auditory-vocal performance and cannot be attributed to a general impairment in dual-task performance, as is clear from the finding that performance on the visual tasks was similar in the three groups.

The present data support and extend the findings of Bar-Haim et al (2004) suggesting that deficient auditory processing is significantly involved in speech selectivity in some children with SM, who may resort to speech avoidance as a consequence of

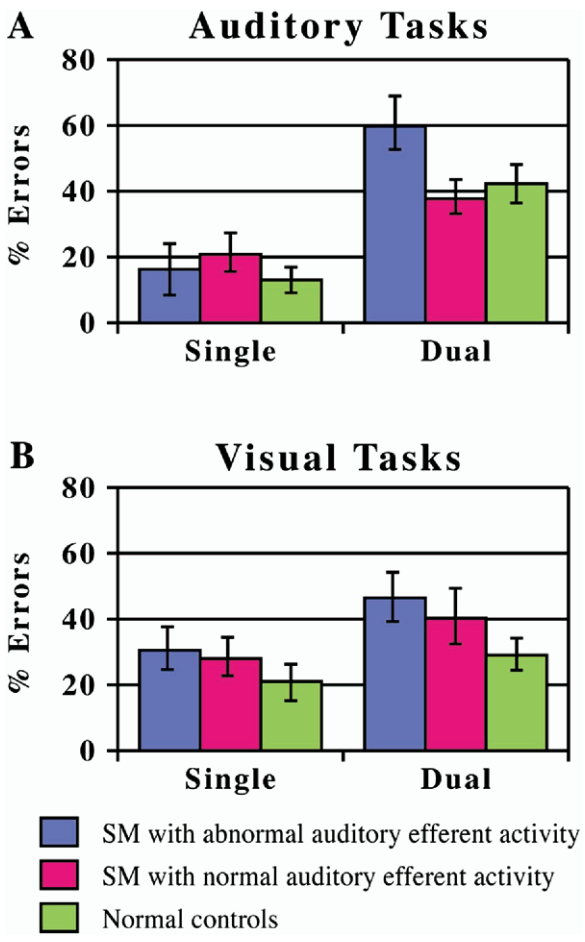

Figure 1. Means and standard deviation bars of percentages of errors in the single and dual auditory tasks (A) and single and dual visual tasks (B), for children with SM with abnormal auditory efferent activity, children with SM with normal auditory efferent activity, and normal control children.

their difficulty in processing incoming sounds while vocalizing. However, it will be important to replicate and extend these findings with larger samples and other experimental tasks, in order to delineate the exact nature of the interplay between efferent dysfunction and speech behavior in SM.

What mechanisms might underlie the selectivity of mutism? Some clues are provided by reports of links between elevated cortisol levels, which have been widely implicated in the etiology and maintenance of anxiety disorders (e.g., Schulkin and Rosen 1999), and increased MEAR thresholds. Fehm-Wolfsdorf et al (1993) showed that following social stress induction participants who responded with elevated cortisol secretion needed significantly higher loudness to elicit the MEAR than participants who did not respond with elevated cortisol secretion. This finding (Fehm-Wolfsdorf and Nagel 1996), along with the findings of Bar-Haim et al (2004), and those of the present study suggest a neurobiological model that may account for both the selectivity of mutism and the high rates of comorbidity between SM and social anxiety (Anstendig 1999; Black and Uhde 1992; Steinhausen and Juzi 1996). We tentatively propose a diathesis-stress model according to which some children with SM are characterized by an auditory neuro-functional vulnerability. For such children, selectivity of speech may be mediated, in specific circumstances, by elevated anxiety that leads in turn to increased cortisol secretion. Increased cortisol levels may in turn interact with the auditory vulnerability to cause elevated MEAR thresholds and other efferent deficiencies. Such stress-induced auditory processing alterations may tax a child's ability to process external sounds during vocalization and in some cases might lead to full blown SM. Further research is needed to establish this diathesis-stress model of SM by measuring cortisol secretion levels and MEAR function before and after induced stress in children with SM.

Because a significant number of children with SM appear to present with elevated MEAR thresholds (2/3 of the Bar-Haim et al 2004 sample and 1/2 of the present sample), it may be adequate to assess MEAR thresholds as part of the diagnostic procedure of SM. MEAR threshold assessment is a noninvasive procedure that does not require vocal cooperation from the child and thus could be completed with children with SM. A diagnosis of this auditory dysfunction can provide a powerful psychoeducational tool for children, parents, and teachers. Reframing a child's mutism in terms of a coping response to a physiological deficit may reflect positively on parent-child and teacher-child relationships. In addition, treatment of auditory processing difficulties might enhance other treatments of SM.

We would like to thank Dr. Liat Kishon-Rabin for her useful advice in the selection of the auditory stimuli, and Inbal Tamir for her invaluable help with data collection.

Supplementary material cited in this article is available online.

Anstendig KD (1999): Is selective mutism an anxiety disorder? Rethinking its DSM-IV classification. *J Anxiety Disord* 13:417–434.
Bar-Haim Y, Henkin Y, Ari-Even-Roth D, Tetin-Schneider S, Hildesheimer M, Muchnik C (2004): Reduced auditory efferent activity in childhood selective mutism. *Biol Psychiatry* 55:1061–1068.
Black B, Uhde TW (1992): Elective mutism as a variant of social phobia. *J Am Acad Child Adolesc Psychiatry* 31:1090–1094.
Borg E, Counter SA, Rosler G (1984): Theories of middle ear muscle function. In: Silman S, editor. *The acoustic reflex: Basic principles and clinical applications*. New York: Academic Press, 63–99.
Borg E, Zakrisson JE (1973): Letter: Stapedius reflex and speech features. *J Acoust Soc Am* 54:525–527.
Borg E, Zakrisson JE (1975): The activity of the stapedius muscle in man during vocalization. *Acta Otolaryngol* 79:325–533.
Dewson JH (1968): Efferent olivocochlear bundle: Some relationships to stimulus discrimination in noise. *J Neurosci* 31:122–130.
Fehm-Wolfsdorf G, Nagel D (1996): Differential effects of glucocorticoids on human auditory perception. *Biol Psychol* 42:117–130.
Fehm-Wolfsdorf G, Soherr U, Arndt R, Kern W, Fehm HL, Nagel D (1993): Auditory reflex thresholds elevated by stress-induced cortisol secretion. *Psychoneuroendocrinology* 18:579–589.
Giraud AL, Garnier S, Micheyl C, Lina G, Chays A, Chery-Croze S (1997): Auditory efferents involved in speech-in-noise intelligibility. *Neuroreport* 8:1779–1783.
Hoy R (2002): Tuning in by tuning off. *Nature* 418:831–833.
Micheyl C, Collet L (1996): Involvement of olivocochlear bundle in the detection of tones in noise. *J Acoust Soc Am*99:1604–1610.
Schulkin J, Rosen JB (1999): Neuroendocrine regulation of fear and anxiety. In: Schmidt LA, Schulkin J, editors. *Extreme fear, shyness, and social phobia: Origins, biological mechanisms, and clinical outcome*. New York: Oxford University Press, 140–172.
Steinhausen HC, Juzi C (1996): Elective mutism: an analysis of 100 cases. *J Am Acad Child Adolesc Psychiatry* 35:606–614.