Out of the Mouths of Babes: Vocal Production in Infant Siblings of Children with ASD

Rhea Paul, Ph.D.¹, Yael Fuerst, M.A.², Gordon Ramsay, Ph.D.¹,³, Kasia Chawarska, Ph.D.¹, and Ami Klin, Ph.D.¹
¹Yale Child Study Center
²Southern Connecticut State University
³Haskins Laboratories

Keywords
autism; speech; vocalization; infant siblings

Autism spectrum disorders (ASD) are developmental disabilities marked by severe deficits in reciprocal social interaction, communication, as well as repetitive and restricted patterns of interests and behavior (APA, 1994; Volkmar, Lord, Bailey, Schultz, & Klin, 2004). Children who qualify for diagnoses of ASD in the second year of life consistently demonstrate significant delay in the development of expressive language (Chawarska, Klin, Paul, & Volkmar, 2007; Chawarska & Volkmar, 2005; Paul, Chawarska, Cicchetti, & Volkmar, 2008; Wetherby, Watt, Morgan, & Shumway, 2006). Other studies of early vocal behavior in infant siblings of children with ASD (Landa & Garrett-Mayer, 2006; Landa, Holman, & Garrett-Mayer, 2007; Iverson & Wozniak, 2007; Zwaigenbaum, Bryson, Rogers, Roberts, Brian, & Szatmari, 2005) have reported both delays on standard tests of expressive language, and less mature syllable production (Dawson et al., 2000). In addition, several studies have shown that preschoolers with ASD produce higher levels of non-speech-like, atypical vocalizations than peers with typical development (Oller et al., 2010; Schoen et al., 2010; Sheinkopf, Mundy, Oller, & Steffens, 2000; Wetherby et al., 2006; Wetherby, Woods, Allen, Cleary, Dickinson & Lord, 2004).

Risk for ASD is known to be elevated in both parents and siblings of affected children (Bailey et al., 1995; Dawson, Estes, Munson, Schellenberg, Bernier & Abbott, 2007; Folstein & Rutter, 1977, 1988; Landa, Piven, Wzorek, Gayle, Chase, & Folstein, 1992; Piven, Palmer, Jacobi, Childress & Arndt, 1997; Santangelo & Folstein, 1999). In addition, Le Couteur, Bailey, Goode, Pickles, Robertson, and Gottesman (1996) found that, in twin pairs discordant for autism, concordance for language impairment and a broader autism phenotype (BAP) was much greater in monozygotic pairs than dizygotic pairs, consistent with a strong genetic component. Although definitions are still evolving, Miceli, Chakrabarti, & Fombonne (2004) present a definition of BAP based on epidemiological findings, which includes social and communication impairments, as well as repetitive and stereotyped behaviors. Social abnormalities have been found to consist of impaired friendships, impaired social play, odd behavior and impaired conversation. Developmental language abnormalities are seen particularly in pragmatic language areas (Landa et al., 2006; Volkmar et al., 2004).

Address for correspondence: Rhea Paul, Director, Communication Disorders Section, Yale Child Study Center, 40 Temple St. # 6B, New Haven, CT 06510, Ph.: 203 785 3388, Fax: 203 785 3705, rhea.paul@yale.edu.
1992). Obsessional and repetitive behaviors can include circumscribed interests, rigidity, obsessions/compulsions and repetitive interests and activities. Micali et al. (2004) define BAP as a constellation of these abnormalities at a subthreshold level of severity that would qualify an individual for neither ASD nor another diagnosis, such as developmental delay or anxiety disorder.

Recent research on ASD has included ongoing monitoring of infants with an older sibling with ASD. This research focus stems from the literature reviewed above, which suggests that genetic factors play a significant role in the incidence of ASD (Bailey et al., 1995; Folstein & Rutter, 1988; Le Couteur et al., 1996; Rutter, 2005). In addition, awareness of risk for disorders associated with BAP, even in the absence of the full-blown syndrome (Szatmari et al., 2000), contributes to this trend. Studies documenting infant siblings’ development are aimed at earlier identification of these potential problems, and earlier provision of treatment in order to maximize outcome potential.

In the present report, we describe findings on early vocal behaviors in a group of infant siblings with this risk, and compare these behaviors to those of age-mates from low-risk families. In doing so, we would like to argue that the direct and detailed analysis of vocal behaviors in infants at risk for ASD will provide richer information than scores on standard tests, such as Mullen (1995), or than retrospective parent report of early vocal behavior. First, these analyses can tell us not only that high risk infants are delayed relative to peers, but also in precisely what aspects of prespeech behavior, whether it be the acquisition of specific consonants, syllable shapes, or prosodic contours. This level of detail may be illuminating in better understanding how the path to spoken language is affected in high risk infants. Second, although parents are sensitive to certain milestones of speech development (first canonical syllables [Oller, Eilers, Neal, & Cobo-Lewis, 1998], first words), they are not generally attuned to the specific consonants their children are using or the relative distribution of consonant and syllable types in infant production. Finally, retrospective parent reports, which have been used in other studies of early speech development (e.g., Gernsbacher, Sauer, Geye, Schweigert, & Goldsmith, 2008) are subject to errors of recall and bias. By directly examining the maturity of prelinguistic vocalizations, we test the hypothesis that infant siblings of children with ASD will show precursors to delays in expressive language that are consistent with the development of ASD or BAP at higher rates than will low risk infants, and we examine the relation of these delays in the first year of life to the presence of autistic symptomatology in the second year.

Vocal Production in Typical Development

Oller, Eilers, Neal, & Schwartz (1999), Stark (1980), Stark, Bernstein, and Demorest (1993), and Locke (1993) have described the course of vocal development in the first year of life. Major milestones include:

- The appearance of first consonant-like productions in vocal play at 4-6 mo.
- The predictable expansion of the consonant repertoire; Shriberg (1993) classified the consonants of English into three groups, based on their typical order of acquisition in the first year (See Table 1).
- The emergence of canonical babble at 6-10 mo.; i.e., the use of consonant-vowel syllables with speech-like voice quality (Oller et al., 1999).

Studies of typically developing children (e.g., Oller, Wieman, Doyle, & Ross, 1976; Schwartz & Leonard, 1982; Stoel-Gammon & Cooper, 1984; Vihman, Ferguson, & Elbert,
1986) have shown that canonical babble in the second half of the first year of life is
generally continuous with the development of first sounds and words in speech, and
babbling behavior at the end of the first year predicts later speech and language production
in children with both typical development (Jensen, Boggild-Andersen, Schmidt, Ankerhus,
& Hansen, 1988; Leonard, Schwartz, Morris, & Chapman, 1981; Menyuk, Liebergott, &
Schultz, 1986; Stoel-Gammon, 1998; Storkel & Morrisette, 2002; Vihman & Greenlee,
1987) and those with delayed language (Paul & Jennings, 1992; Roberts, Rescorla, Giroux,
& Stevens, 1998; Thal, Oroz, & McCaw, 1995). Thus investigating the development of pre-
speech vocalization in the latter half of the first year may provide some insight into the
transition to language in this population.

In the present report, we examine the forms of both speech-like and non-speech
vocalizations in samples of infants at high and low risk for ASD, seen at 6, 9, and 12 months
of age, using a cross-sectional design. In this way we aim to investigate the vocal tools
children at risk for ASD are bringing to the task of learning to communicate. The following
hypotheses will be tested:

1. HR infants will produce fewer consonants and canonical syllable shapes, and more
   non-speech vocal behavior than LR peers;

2. Less frequent and mature prespeech vocal behavior in HR infants during the first
   year of life will be associated with greater likelihood of the appearance of autistic
   symptomatology in the second year.

**METHOD**

**Participants**

Infants were seen as part of a longitudinal study of the development of children at high and
low risk for ASD. The study was approved by the Institutional Review Board of the Yale
School of Medicine for human subject research, and all families signed informed consents
prior to participation. Participants were assigned to the High Risk (HR) group if a full
sibling had received a diagnosis of ASD by an experienced professional or multidisciplinary
team. Many of the proband siblings’ families had participated in research on toddlers with
ASD at our center, had been diagnosed by our team and were engaged in ongoing follow-up.
Other proband older siblings’ families were referred through ongoing relationships with
local pediatric practices, local autism advocacy organizations, and word of mouth from
families already participating in our HR infant follow-up. Diagnoses of older siblings were
evaluated by review of clinical reports of previous evaluations and confirmed by means of
the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & LeConte, 1994)
conducted with parents by trained clinicians. Infants were assigned to the Low Risk (LR)
group if no sibling in the family had received a diagnosis of autism or any other
developmental disorder. Families for this group were recruited from local pediatric
practices, advertisements in parenting media, and word of mouth.

Exclusionary criteria for both HR and LR groups included less than 32 weeks gestation,
known sensory deficits such as significant vision or hearing loss, known genetic syndrome,
or known neurological or significant chronic medical disorder.

For the longitudinal study from which the present study draws data, families are invited for
multiple visits during the infants’ first year and for follow-up at 24 and 36 mo. During these
visits, a variety of experimental procedures on visual and auditory processing were
administered. From six months onward, visits also included clinical characterization
measures.
Participants in the present report include those who participated in vocalization sample collection at the 6 (HR: n = 28; LR: n = 20), 9 (HR: n = 37; LR: n = 29), and/or 12 (HR: n = 38; LR: n = 31) month visit. In addition, a subgroup of these who had contributed vocalization samples have by now reached age two and participated in a 24 month follow-up (HR: n=24; LR: n = 21). The remaining children had not reached 24 months at the time of this writing. It should be noted that although some subjects participated in more than one visit, not all families attended every visit. The number of families who participated in all four visits (n=14) was too small to allow for treating the sample longitudinally; therefore data were analyzed cross-sectionally.

Participants who attended the 24 month visit received a consensus provisional diagnosis based on clinical observations and standard assessments conducted by experienced clinicians blind to the participants’ risk status. Clinicians examined all information collected at the visit and conferred one of the following consensus provisional diagnoses: ASD, Broad Autism Phenotype (BAP), Non-autistic Developmental Delay (DD), or No Clinical Diagnosis (i.e., typical development).

Procedures

**Standard Measures**—Standard characterization measures included the *Autism Diagnostic Observation Scale*-Toddler Module (ADOS-T; Lord et al., 2002; Luyster, Guthrie, Gortham, Risi, DiLavore, & Lord, 2009) and the *Mullen Scales of Early Learning* (MSEL; Mullen, 1995). It should be noted that at each visit at which the MSEL was administered (6, 12, and 24 mo.), there was no significant difference between the HR and LR groups in terms of the Visual Reception scale. Although the HR group’s scores tended to be lower, as Table 2 shows, the groups were statistically equivalent in terms of this measure of non-verbal cognition.

**Provisional Clinical Diagnosis**—Provisional diagnoses were conferred at the 24 month visit by consensus judgment from two highly experienced clinicians. The average ADOS-T total scores for the HR children who received diagnoses of ASD, BAP, or No ASD Symptoms appear in Figure 1. Clinical diagnoses were always arrived at by consensus among experienced clinicians.

**ASD:** Provisional diagnosis of ASD was based Chawarska et al.’s (2009) model employing data both from the ADOS-T and the pattern of *MSEL* (1995) verbal and non-verbal scores, as well as clinical observation and judgment. HR children receiving this diagnosis had a significant number of points on the ADOS-T (at least 5), had a pattern of verbal and non-verbal performance on the MSEL suggestive of that seen in toddlers with ASD (i.e., poorer verbal than non-verbal skills, poorer receptive than expressive language scores on the MSEL), and qualitatively demonstrated more severe, more atypical, and more consistently aberrant social behavior than children in the other diagnostic groups.

**BAP:** HR children receiving a provisional diagnosis of BAP also evidenced some elevation on ADOS-T scores (all those diagnosed with BAP had ADOS-T scores of at least 8). However, the abnormalities in their social interactions, the severity of their atypicality and the consistency of their aberrant behaviors were less pronounced than those diagnosed with ASD. Clinical judgment based on extensive observation and review of parental report was used to assess this. Children assigned to the BAP category also did not exhibit the characteristic pattern of MSEL scores seen in ASD (i.e., Visual Reception>Express Language>Receptive Language). Although some children who were diagnosed as BAP had atypical language behaviors, such as echolalia or scripting, none had language delay alone, in the absence of other social difficulties, nor did any language delays they showed qualify...
them for a diagnosis of Delayed Language development, according to Connecticut Birth-to-
Three criteria (i.e., more than one standard deviation below the mean in both expressive and
receptive language, or more than two s.d. below the mean in either one).

**No ASD Symptoms:** Children received no diagnosis (i.e., typical development) if their
social behavior was judged to be within the typical range for age (ADOS-T scores ranged
from 3-9), and they showed no significant delays on the MSEL. Developmental Delay (DD)
was conferred when a child scored more than one standard deviation below average on at
least one scale of the MSEL, but showed no significant social deficits on observation or by
parent report. Only one HR child in this cohort was assigned to this category.

**Vocalization Sample Collection**—Samples were collected from a timed parent-child
interaction. Parents were asked to play and interact with the child with a standard set of toys
for five minutes, with no warm-up period. Vocalization samples were recorded using a
Shure SM58 omni-directional microphones placed in a stand on the floor close to the child.
A Marantz CDR300 compact disk recorder captured the vocalizations on digital media.

**Transcription**—Infant vocalizations were separated into two categories, speech-like or
non-speech. The speech-like category included vocalizations characterized by the
production of consonants and/or vowels that could be represented by phonetic symbols and
contained speech-like vocal quality such that they resembled typical babble. The non-speech
category included vocalizations characterized by the production of non-speech resonance
and vocal quality (e.g., squeal, yells, growls) without recognizable consonants.

The first 50 speech-like vocalizations produced were transcribed using broad phonemic
transcription. For participants who did not produce 50 speech-like vocalizations within the
time frame, all speech-like vocalizations produced were transcribed. It should be noted that
at 6 mo. approximately 95% of participants in both risk groups produced fewer than 50
speech-like vocalizations; at 9 months this figure was approximately 90%; and at 12 months,
approximately 80%. The same five-minute segments transcribed for speech-like
vocalizations were also coded for non-speech vocalizations. Vegetative sounds including
coughing or burping were not included in analyses of either speech-like or non-speech
vocalizations.

**Coding**—Audiofiles were edited so that speech-like and non-speech productions were
copied onto separate files and stored. Coders were blind to participants’ age and risk group
assignment.

**Frequency of Vocalization:** Frequency of speech-like vocalizations (up to 50) was tallied
for each participant. All non-speech vocalizations occurring within the same time period as
the speech-like sample were tallied. Non-speech vocalizations were separated into utterances
based on breath groups or a pause of greater than one second. Total number of vocal
productions for each participant was derived by summing the number of speech-like and
non-speech vocalizations. The two vocalization types were then coded separately. Average
frequencies of each vocalization type for each risk group at each age appear in Table 3.

**Coding of Speech-like Vocalizations**—Any utterance that could not be confidently
transcribed after four playbacks was eliminated. The following coding was completed for
each participant:

**Consonant Inventory:** Inventories of singleton consonants were assembled for each
participant. Consonants were then divided into three categories as outlined by Shriberg
The total number of consonant types and the number of consonant types in each of the three developmental categories was computed for each participant.

**Percent Canonical Syllables:** The number of syllables containing the basic Consonant-Vowel (CV) structure associated with true babbling (Oller et al., 1999) was tallied for each participant and divided by the total number of speech-like vocalizations to derive this metric.

**Non-speech Coding**

**Proportion of Non-speech Production:** The percentage of non-speech vocalizations was computed for each participant by counting the number of non-speech vocalizations and dividing by the sum of the number of speech-like plus non-speech vocalizations produced. The following categories of non-speech vocalizations were used to identify non-speech productions (Sheinkopf et al., 2000):

1. Delight: Laughing or giggling.
2. Distress: Crying, whining or fussing.
3. Atypical: High-pitched squeals, low-pitched growls, yells, grunts.

**Reliability—**Two trained raters, blind to subject age and risk status, independently scored a 10% randomly selected sample of the vocalizations on all coding categories. Point-to-point reliability exceeded 80% on all coding categories.

**Analytic Strategy**

*MSEL* scores were examined at each age-level for which data had been collected (6, 12, and 24 mo.), and one-way analysis of variance at each age was carried out to compare scores on the five *MSEL* subtests between the two risk groups. These data appear in Table 2. Two-way analysis of variance, with age (6, 9, 12 mo.) and risk group (HR, LR) as independent variables, was used to analyze data on each of the eight vocalization behaviors studied (total # vocalizations, # speech-like vocalizations, # consonants, # early consonants, # middle consonants, # late consonants, % canonical syllables, % non-speech vocalizations). Main effects for age and risk group, as well as the interaction of age and risk group were examined. In cases where significant interaction effects were found, pair-wise comparisons between risk groups at each age level (6, 9, and 12 mo.) were conducted, and significant pair-wise differences, along with effect sizes (Cohen, 1988) are reported. Data input to these analyses are presented in Table 3.

Correlations between *MSEL* scores and vocal productions were also examined within the HR group at the two ages for which *MSEL* scores were available (6 and 12 mo.). Discriminant function analysis was then used to predict placement in groups with and without autistic symptoms at a 24 month follow-up visit for the HR group only (since prediction of outcome for HR children was the question of interest), using the categories of vocal production on which risk group differences were found in the first year data.

**RESULTS**

**Total Frequency of Vocalizations**

ANOVA indicated that there was a main effect for age (F(2)=3.8; p=.02), but not for risk, and no significant age × risk interaction for total number of vocalizations. These findings suggest that all infants produced more vocalizations as they got older, but there was no significant difference between risk groups in frequency of vocal production, when both
speech and non-speech vocalizations were combined. Thus, differences between groups cannot be attributed to less frequent vocal behavior on the part of the HR group.

**Frequency of Speech-like Vocalizations**

There was a significant main effect for age (F(2)=8.3; p=.0003) and risk group (F(1)=7.9; p=.006) but no significant interaction for frequency of speech-like vocalization. This main effect suggests that speech-like productions increased with age for all infants, and the HR children produced less speech-like vocalization at all ages.

**Percent Canonical Syllables**

There were significant main effects for age (F(2)=29.6; p<.0001) and risk group (F(1)=6.0; p=.001) and a significant age x risk interaction in terms of the proportion of speech-like productions that consisted of canonical syllables. Main effects suggest that canonical syllable production increased with age overall, and that there is a difference in canonical syllable production between risk groups. Exploring the interaction effect with post-hoc pair-wise comparisons, using the Bonferroni correction reveals a significant difference between risk groups at the 9 month level only (p=.04; Cohen’s d=0.79) with a medium to large effect size.

**Consonant inventories**

Significant main effects were found for both age (F (2)=20.1; p<.0001) and risk group (F (1)=11.4; p=.0009) in terms of the total number of consonants produced. All infants produced more consonants as they got older, and those in the LR group produced more consonants at each age than those in the HR group.

**Consonant classes**—To look in more detail at the production of consonants in this sample, we divided the consonants produced into the three developmental classes presented by Shriberg (1993; See Table 1).

Results for each of the three classes were similar. In each case, there were significant main effects for age (Early C: F(2)=24.1; p<.0001; Middle C: F(2)=11.2; p<.0001; Late C: F(2)=3.7; p=.03) and risk group (Early C: F(1)=8.7; p=.004; Middle C: F(1)=9.0; p=.003; Late C: F(1)=5.4; p=.02). The interaction effect for late consonants approached significance at p<.09. Pair-wise comparisons revealed significant differences between groups at 9 months (p<.03; d=. 78, large) on this variable.

**Non-speech Vocalizations**

The proportion of non-speech production out of total vocalizations also showed main effects for age (F(2)=20.1; p<.0001) and risk status (F(1)=11.0; p=.001) with no significant interaction. Both groups decreased the proportion of non-speech production with age, but the HR infants tended to produce more non-speech at each age.

**Correlations between MSEL Scores and Vocal Production in HR Infants**

At 6 months there were no significant correlations between any of the MSEL scores and vocal production data, suggesting that vocal production is somewhat independent of developmental level at this age. At 12 months, there were no significant correlations between MSEL Gross Motor, Visual Reception, or Receptive Language scores and any of the vocalization variables. There was a significant correlation, however, between the number of Early consonants produced and both MSEL Fine Motor (r=.43, p=.007) and Expressive Language (r=.49, p=.002) scores. Early consonants were still the most frequent consonant class at 12 months, and the correlations between these two MSEL scores and the total
number of consonants produced at 12 months was also marginally significant (r=.38, p=.018; and r=.40, p=.014, respectively). No other correlations reached significance at p<.01.

Follow-up Data at 24 months

Provisional Diagnostic Assignments—As of this writing, 24 HR and 21 LR participants from whom vocalization data were collected have reached their second birthday and have been seen for follow-up evaluations at 24 months of age. One child of the 21 in the LR risk who participated in the 24 month visit was given a provisional ASD diagnosis; 2 others received provisional diagnoses of other, Non-autistic Developmental Delay (it should be noted that parents of LR children who chose to participate in this study may have done so because of underlying concerns about their child’s development or behavior; 7% of parents of LR children who enrolled in the study at 6 mo. of age indicated that they had some developmental concerns about their infants). In the HR group, 7 (29%) received a provisional diagnosis of ASD; 6 (25%) others were seen to exhibit some symptoms without meeting full clinical criteria, for a provisional diagnosis of BAP; 1 (4%) showed Non-autistic Developmental Delay; and ten (42%) were considered to merit no clinical diagnosis. Discriminant function analyses included only HR children, since this is the group for which we aimed to discriminate outcomes on the basis of early vocal data.

It should be noted that diagnoses for high risk infants at 24 months must be considered provisional. Although several research groups (e.g., Chawarska et al., 2009; Cox et al., 1999; Lord, 1995; Turner and Stone, 2007) have reported relative stability of ASD diagnoses in two-year olds in the general population, HR infants show more varying symptoms (Ozonoff et al., 2009; Rogers, 2009). Thus the fact that nearly a third of HR infants appear to have a diagnosis of ASD at 24 months, and another quarter appear to show BAP does not mean that these diagnoses will inevitably remain stable. Moreover, this figure is not unlike findings of other studies of HR siblings seen for assessment at 24 months. For example, both Zwaigenbaum et al. (2005), and Sullivan et al. (2007) found that 30% of an HR sample met criteria for ASD at 24 months.

Discriminating Outcomes in Terms of Autistic Symptoms—In order to determine whether any aspects of prelinguistic vocal development were associated with the appearance of autistic symptoms at the 24 month visit, discriminant function analyses were performed. Participants in the HR group who took part in the 24 month visit were subdivided into two groups: those with autistic symptoms (n=14); i.e., those to whom clinicians assigned a provisional diagnosis of either ASD or BAP; and those who received no diagnosis or a diagnosis of Non-autistic Developmental Delay (n=11). The seven vocal measures on which risk group differences were found (# speech-like vocalizations, # consonants, # early consonants, # middle consonants, # late consonants, % canonical syllables, % non-speech vocalizations) were entered step-wise into each of the discriminant function analysis.

Classification Based on 6-Month Vocal Data: Discriminant function analysis (SPSS 16.0) revealed that only the number of middle consonant types produced by the 17 HR children seen at both six months and 24 months was significant in classifying participants with and without autistic symptoms at 24 mo., with a canonical correlation of .47 (p<.04). Seventy-four percent of subjects were correctly classified on the basis of this variable. One-way ANOVA comparing the 6 mo. MSEL scores of the children who did and did not show symptoms of ASD at 24 months revealed a significant difference (F(1, 21); p<.05; d=.88) in Expressive Language scores only, with children who later showed no symptoms scoring higher (Mean: 41.1; s.d., 3.8) than those who developed symptoms (Mean: 36.4, s.d. 4.5).
Classification Based on 9-Month Vocal Data: Entering the same vocal measures collected at the nine-month visit in a step-wise fashion to predict placement for the 19 HR children seen at both 9 months and at 24 months revealed that only the number of late consonant types produced at nine months was significant in classifying participants (canonical correlation = .53; p < .03), with 77% of subjects were correctly classified on the basis of this variable. No MSEL data are available at this age.

Classification Based on 12-month Vocal Data: Finally, entering the same vocal measures collected for the 20 HR children seen at both 12 and 24 months in a step-wise fashion to predict placement in the group with or without autistic symptoms at 24 months, discriminant function analysis revealed that only total number of different consonant types produced was significant in classifying participants (canonical correlation = .43; p < .03), with 65% of subjects correctly classified on the basis of this variable. One-way ANOVA comparing the 12 mo. MSEL scores of the children who did and did not show symptoms of ASD at 24 months revealed no significant differences on any of the MSEL scales.

Discussion

Apart from showing that all infants progress with age in all the prespeech variables studied, these data also allow us to address the hypotheses stated at the outset. First, it does appear that infants at risk for ASD show differences in some prelinguistic vocal behaviors relative to LR peers; specifically, we found that:

• On average, HR infants produced significantly fewer speech-like vocalizations and more non-speech vocalization than LR peers;
• On average, HR children produced significantly fewer consonant types than LR peers;
• HR infants produced significantly fewer canonical syllable shapes than LR peers, particularly at 9 months of age.

Second, differences in vocal production in the first year of life were associated with outcomes in terms of autistic symptomatology in the second year for the HR group. Differences seen in prespeech vocal production of HR infants cannot be ascribed to a general decrement in developmental level, as their MSEL Visual Reception scores do not differ from those of LR peers at any time point measured. Rather, they appear to reflect a specific delay in acquiring age-appropriate sounds and syllable shapes in the first year, and in transitioning from the diverse range of sound production typical of all the young babies to a more restricted range dominated by speech-like sounds at the first year’s end. These delays are captured not only by the detailed analysis of sound production reported here; they are hinted at by MSEL Expressive Language scores, which are significantly lower than those of the LR group throughout the first year, although these scores are within the normal range. Moreover, Expressive Language scores on the MSEL correlate with vocal production of consonants at 12 months. This is likely a reflection of the fact that most items used to score Expressive Language on the MSEL at the 12 month level relate to consonant production. It is also noteworthy, however, that fine motor skills are also related to consonant production at this age.

These findings can be interpreted to suggest that pre-speech sound development is a sensitive indicator of the rate and degree to which infants at risk for ASD are following the developmental path to language acquisition, and, by extension, to expressive communication. Motor delays cannot be ruled out as a source of the pattern seen here, especially in light of the correlation between Fine Motor scores on the MSEL and consonant
production at 12 months. Despite this correlation, though, findings from the MSEL suggest that, apart from an apparent mild delay in Fine Motor performance at 6 months, no significant motor delays persist into the second half of the first year.

Does this detailed analysis of vocal production in infants at risk for ASD contribute any information beyond what is measured by the more general MSEL language scales? Potential answers to this question may be seen in the changing pattern of relationship between the vocal performance of the HR and LR groups throughout the first year of life, as well as the findings of the discriminant function analyses. That is, detailed description of vocal production reveals that the areas of difference consistently reflect slower progress on the part of the HR infants in moving from one phase of pre-speech development to the next. By 9 months -- when LR children are developing a range of tools necessary to begin the transition from pre-speech to first speech production -- HR children produce, on average, a lower rate of production of canonical syllables, the kind that will eventually be used as first words, and a smaller repertoire of consonants, particularly the Late group, those just emerging at this developmental level. By 12 months, the participants with HR, as a group, are not significantly different from LR peers in production of the consonant and syllable types their LR peers learned some months earlier (although they continue to trend toward less frequent production) but they continue, as they have throughout the first year, to produce a higher proportion of other kinds of vocalizations that do not contribute to their transition to first words (See Fig. 5). As we have argued previously (Schoen et al., 2008) LR infants, tend, toward the end of the first year of life, to match their productions more and more closely to ambient speech models. HR infants, on the other hand, show less winnowing of the other kinds of sounds characteristic of the first year of life. Thus the pattern of pre-speech development seen in the HR group suggests both slower transition from one phase of pre-speech development to the next, and a reduced tendency to hone sound production increasingly closely to models produced by others in the environment as the child nears the first birthday. This more detailed examination suggests that low MSEL Expressive Language scores have different sources at different ages: slow acquisition of pre-speech behaviors during the first year, and a slower transition away from non-speech vocal production near the first birthday. Moreover, although MSEL Expressive Language scores at 6 mo. appear to be somewhat sensitive to differences between children who will and will not go on to show ASD symptoms at 24 mo., the same is not true at 12 mo., even though MSEL Expressive Language is generally correlated with consonant production at this age. It would appear that although the MSEL Expressive Language score captures some of the decreased prespeech behavior present at 6 mo. in HR infants who go on to develop symptoms, it is less sensitive to the transitional behaviors at 12 months that are associated with later symptom appearance.

It is important to recall that the high prevalence of autistic symptoms in the HR does not necessarily mean that all symptomatic children with remain on the autism spectrum. Thus HR/LR group differences reported here may be driven by the children who have significant vulnerabilities for ASD. The discriminant function analyses are telling in this regard. At each age level, the metric most strongly associated with placement in the outcome group with or without autistic symptoms is the production of newly emerging speech behaviors. That is, at six months it is the frequency of middle consonant production, the consonants just emerging at this age, that predicts group placement; at 9 months, it is the frequency of late consonants, again those at the leading edge of prespeech development. At 12 months, even though group differences in consonant production are no longer significant, it is the total number of consonants, which can be interpreted to reflect the amount of word-like babble, that is associated with outcome.

What might this pattern suggest about spoken language development in children at risk for ASD? It would seem, first, to indicate that it is not only the development of communicative...
intention and language function that is vulnerable in the HR population; the acquisition of
basic vocal “tools” for the development of speech is also affected, even before they are
recruited into the service of linguistic communication. Prior to the onset of language,
babbling is used in typical development both as a means of “practicing” the motor patterns
associated with speech production and as an early form of interaction, through back-and-
forth babbled “conversations” with family members (Locke, 1989). Perhaps vocal
development is vulnerable in HR infants because they are less apt to “practice” sound-
making in order to match it more closely to the sounds they hear others in their environment
producing, as well as because they are less captivated by the turn-taking exchanges in which
babbled productions allow them to participate. Thus, the differences observed here may be a
result of vulnerabilities for social disability, and not their source. It is not the case, however,
that HR infants are simply “quieter” than LR peers, since the total number of vocalizations
produced never differed between the two risk groups at any age. Differences appear not in
sound production, but in the kinds of sounds produced, with LR children and HR children
who do not develop ASD symptoms producing more vocal behavior at the “cutting edge” of
development at each phase.

Limitations

Although delayed language development is a core feature of ASD, it is not specific to ASD,
and the lack of a non-autistic developmentally delayed contrast group is a limitation of the
present study. It could be the case that the differences in preverbal behavior in the HR
infants are associated with a more general or more language-specific delay than with ASD.
Although our discriminant function analysis argues for a connection between preverbal
vocal behavior and autistic symptoms rather than general delay, and ANOVA results
revealed there were no differences in either nonverbal or receptive language skills at any age
between HR children who did and did not develop ASD symptoms by 24 months,
nonetheless, a direct contrast of preverbal vocal production in children with non-autistic
developmental delays in the first year of life and those at risk for ASD is clearly warranted.
In addition, a true longitudinal design would provide a more direct answer to the questions
raised in this study, which was limited to a cross-sectional design.

Clinical Implications

Although the limitations of the present study prevent drawing a definitive clinical
connection between early speech behavior and symptoms of ASD in HR infants, the
findings do suggest that children showing delays in speech acquisition, whether measured by
direct observations of consonant inventories and canonical syllables, or more indirect
measures such as the MSEL Expressive Language scale, warrant extra vigilance during the
second half of their first year. This vigilance may include more frequent monitoring of a
range of developmental milestones. In addition, parent counseling that guides families to
maximize the infant’s experience in listening to and making speech sounds, such as that
proposed by Kuhl & Meltzoff (1996), may be considered along with interventions that
encourage the development of joint attention and communicative intent. The finding that
pre-speech vocal development is a sensitive indicator of heightened risk for the appearance
of ASD symptoms in HR infants provides an addition to the range of tools emerging from
the infant sibling literature in ASD for the important task of identifying children at the
earliest possible point in development. Early identification, combined with advice to
families that helps intensify and focus their natural interactions with these infants on
activities aimed at optimizing early interpersonal experiences may allow us alter their
developmental trajectory for the better.
Acknowledgments

Preparation of this paper was supported by NIMH Autism Center of Excellence grant P50 MH81756; Research Grant P01-03008 funded by the National Institute of Mental Health (NIMH); MidCareer Development Award K24 HD045576 funded by NICD to Rhea Paul; as well as by the Simons Foundation. We wish to thank the families who participated in this research.

References


J Child Psychol Psychiatry. Author manuscript; available in PMC 2012 May 1.


Shriberg L. Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorder. Journal of Speech and Hearing Disorders. 1993; 36:105–140.


Figure 1.
ADOS-T total algorithm scores for three groups of HR infants at 24 months. *Bars represent one standard deviation around the group mean.
Table 1


<table>
<thead>
<tr>
<th>Early 8:</th>
<th>/m/</th>
<th>/b/</th>
<th>/j/</th>
<th>/n/</th>
<th>/w/</th>
<th>/d/</th>
<th>/p/</th>
<th>/h/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle 8:</td>
<td></td>
<td></td>
<td>/θ/</td>
<td>/s/</td>
<td>/z/</td>
<td>/ð/</td>
<td>/ł/</td>
<td>/r/</td>
</tr>
<tr>
<td>Late 8:</td>
<td>/∫/</td>
<td>/ð/</td>
<td>/θ/</td>
<td>/s/</td>
<td>/z/</td>
<td>/ð/</td>
<td>/ł/</td>
<td>/r/</td>
</tr>
</tbody>
</table>

* Based on Stoel-Gammon (1998); Locke (1989; 1993).
### Table 2
Mean (and s.d.) MSEL\(^{1}\) T-Scores at Three Visits for High- Vs. Low-Risk Groups

<table>
<thead>
<tr>
<th></th>
<th>VR(^{2})</th>
<th>FM(^{3})</th>
<th>GM(^{4})</th>
<th>RL(^{5})</th>
<th>VR(^{2})</th>
<th>FM(^{3})</th>
<th>GM(^{4})</th>
<th>RL(^{5})</th>
<th>EL(^{6})</th>
<th>VR(^{2})</th>
<th>FM(^{3})</th>
<th>GM(^{4})</th>
<th>RL(^{5})</th>
<th>EL(^{6})</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR</td>
<td>43.5 (8.7)</td>
<td>40.1 (7.9)</td>
<td>44.2 (8.7)</td>
<td>43.6 (9.9)</td>
<td>39.0 (5.3)</td>
<td>53.4 (10.4)</td>
<td>57.0 (10.6)</td>
<td>45.5 (11.3)</td>
<td>41.1 (9.7)</td>
<td>40.8 (10.7)</td>
<td>54.9 (12.6)</td>
<td>52.9 (11.9)</td>
<td>47.3 (4.9)</td>
<td>53.0 (9.8)</td>
</tr>
<tr>
<td>LR</td>
<td>46.4 (8.6)</td>
<td>45.4 (7.2)</td>
<td>44.8 (7.2)</td>
<td>46.5 (8.2)</td>
<td>42.2 (5.8)</td>
<td>58.1 (9.4)</td>
<td>56.3 (10.7)</td>
<td>45.0 (14.0)</td>
<td>46.1 (7.2)</td>
<td>47.4 (12.0)</td>
<td>59.7 (11.4)</td>
<td>54.9 (6.9)</td>
<td>49.5 (7.5)</td>
<td>61.6 (10.2)</td>
</tr>
</tbody>
</table>

HR = High Risk group, LR = Low Risk group

\(^{1}\) Mullen Scales of Early Learning (Mullen, 1995) T scores (mean 50; s.d., 10)

\(^{2}\) Visual Reception (non-verbal problem solving) Scale

\(^{3}\) Fine Motor Scale

\(^{4}\) Gross Motor Scale

\(^{5}\) Receptive Language Scale

\(^{6}\) Expressive Language Scale

* Significant difference between High and Low Risk groups at p<.05
# Table 3

Mean (and s.d.) Vocal Productions in Two Risk Groups

<table>
<thead>
<tr>
<th>Age</th>
<th>Risk Group</th>
<th>Total Vocalizations (Speech + Non-speech)</th>
<th># Speech-like Vocalizations</th>
<th>Total # Consonants</th>
<th># Early Consonants</th>
<th># Middle Consonants</th>
<th># Late Consonants</th>
<th>% Canonical Syllables</th>
<th>% Non-speech Productions</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 mo.</td>
<td>HR</td>
<td>36.5 (27.4)</td>
<td>15.0 (160)</td>
<td>2.5 (29)</td>
<td>1.3 (1.7)</td>
<td>0.5 (1.1)</td>
<td>0.5 (0.9)</td>
<td>4.5 (7.0)</td>
<td>60.2 (28.6)</td>
</tr>
<tr>
<td></td>
<td>LR</td>
<td>39.2 (27.8)</td>
<td>18.3 (153)</td>
<td>3.7 (39)</td>
<td>2.1 (2.2)</td>
<td>1.1 (1.4)</td>
<td>0.5 (0.9)</td>
<td>8.3 (115)</td>
<td>57.0 (25.2)</td>
</tr>
<tr>
<td>9</td>
<td>HR</td>
<td>25.7 (19.6)</td>
<td>15.1 (162)</td>
<td>2.7 (30)</td>
<td>1.9 (1.7)</td>
<td>0.5 (0.9)</td>
<td>0.4 (0.8)</td>
<td>8.7 (115)</td>
<td>50.6 (17.7)</td>
</tr>
<tr>
<td></td>
<td>LR</td>
<td>31.1 (18.6)</td>
<td>22.6 (164)</td>
<td>5.6 (39)</td>
<td>3.2 (2.1)</td>
<td>1.3 (1.4)</td>
<td>1.1 (1.1)</td>
<td>20.5 (17.6)</td>
<td>34.0 (24.2)</td>
</tr>
<tr>
<td>12</td>
<td>HR</td>
<td>35.2 (18.1)</td>
<td>12.6 (168)</td>
<td>6.2 (3.6)</td>
<td>3.8 (2.2)</td>
<td>1.6 (1.2)</td>
<td>0.8 (0.9)</td>
<td>27.1 (229)</td>
<td>36.5 (28.0)</td>
</tr>
<tr>
<td></td>
<td>LR</td>
<td>39.1 (18.1)</td>
<td>31.9 (157)</td>
<td>7.3 (37)</td>
<td>4.5 (2.1)</td>
<td>1.8 (1.5)</td>
<td>1.0 (1.1)</td>
<td>28.8 (190)</td>
<td>19.3 (16.1)</td>
</tr>
</tbody>
</table>

HR = High Risk group; LR = Low Risk group

* Significant difference between groups at p < .05