

# Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development

D. K. Oller<sup>a,b,1</sup>, P. Niyogi<sup>c</sup>, S. Gray<sup>d</sup>, J. A. Richards<sup>d</sup>, J. Gilkerson<sup>d</sup>, D. Xu<sup>d</sup>, U. Yapanel<sup>d</sup>, and S. F. Warren<sup>e</sup>

<sup>a</sup>School of Audiology and Speech-Language Pathology, University of Memphis, Memphis, TN 38105; <sup>b</sup>Konrad Lorenz Institute for Evolution and Cognition Research, Altenberg, Austria A-3422; <sup>c</sup>Departments of Computer Science and Statistics, University of Chicago, Chicago, IL 60637; <sup>d</sup>LENA Foundation, Boulder, CO 80301; and <sup>e</sup>Department of Applied Behavioral Science and Institute for Life Span Studies, University of Kansas, Lawrence, KS 66045

Edited\* by E. Anne Cutler, Max Planck Institute for Psycholinguistics, Heilig Landstichting, Netherlands, and approved June 21, 2010 (received for review April 21, 2010)

For generations the study of vocal development and its role in language has been conducted laboriously, with human transcribers and analysts coding and taking measurements from small recorded samples. Our research illustrates a method to obtain measures of early speech development through automated analysis of massive quantities of day-long audio recordings collected naturalistically in children's homes. A primary goal is to provide insights into the development of infant control over infrastructural characteristics of speech through large-scale statistical analysis of strategically selected acoustic parameters. In pursuit of this goal we have discovered that the first automated approach we implemented is not only able to track children's development on acoustic parameters known to play key roles in speech, but also is able to differentiate vocalizations from typically developing children and children with autism or language delay. The method is totally automated, with no human intervention, allowing efficient sampling and analysis at unprecedented scales. The work shows the potential to fundamentally enhance research in vocal development and to add a fully objective measure to the battery used to detect speech-related disorders in early childhood. Thus, automated analysis should soon be able to contribute to screening and diagnosis procedures for early disorders, and more generally, the findings suggest fundamental methods for the study of language in natural environments.

vocal development | automated identification of language disorders | all-day recording | automated speaker labeling | autism identification

Typically developing children in natural environments acquire a linguistic system of remarkable complexity, a fact that has attracted considerable scientific attention (1–6). Yet research on vocal development and related disorders has been hampered by the requirement for human transcribers and analysts to take laborious measurements from small recorded samples (7–11). Consequently, not only has the scientific study of vocal development been limited, but the potential for research to help identify vocal disorders early in life has, with a few exceptions (12–14), been severely restricted.

We begin, then, with the question of whether it is possible to construct a fully automated system (i.e., an algorithm) that uses nothing but acoustic data from the child's "natural" linguistic environment, isolates child vocalizations, extracts significant features, and assesses vocalizations in a useful way for predicting level of development and detection of developmental anomalies. There are major challenges for such a procedure. First, with recordings in a naturalistic environment, isolating child vocalizations from extraneous acoustic input (e.g., parents, other children, background noise, television) is vexingly difficult. Second, even if the child's voice can be isolated, there remains the challenge of determining features predictive of linguistic development. We describe procedures permitting massive-scale recording and use the data collected to show it is possible to track linguistic development with totally automated procedures and to differentiate recordings from typically developing children and children with language-related

disorders. We interpret this as a proof of concept that automated analysis can now be included in the tool repertoire for research on vocal development and should be expected soon to yield further scientific benefits with clinical implications.

Naturalistic sampling on the desired scale capitalizes on a battery-powered all-day recorder weighing 70 g that can be snapped into a chest pocket of children's clothing, from which it records in completely natural environments (*SI Appendix*, "Recording Device"). Recordings with this device have been collected since 2006 in three subsamples: typically developing language, language delay, and autism (Fig. 1*A* and *SI Appendix*, "Participant Groups, Recording and Assessment Procedures"). In both phase I (2006–2008) and phase II (2009) studies (*SI Appendix*, Table S2), parents responded to advertisements and indicated if their children had been diagnosed with autism or language delay. Children in phase I with reported diagnosis of language delay were also evaluated by a speech-language clinician used by our project. Parents of children with language delay in phase II and parents of children with autism in both phases supplied documentation from the diagnosing clinicians, who were independent of the research. Parent-based assessments obtained concurrently with recordings (*SI Appendix*, Table S3) confirmed sharp group differences concordant with diagnoses; the autism sample had sociopsychiatric profiles similar to those of published autism samples, and low language scores occurred in autism and language delay samples (Fig. 1*B* and *SI Appendix*, Figs. S2–S4 and "Analyses Indicating Appropriate Characteristics of the Participant Groups").

Demographic information (Fig. 1*B* and *SI Appendix*, Tables S3 and S4) was collected through both phases: boys appeared disproportionately in the disordered groups, mother's educational level (a proxy for socioeconomic status) was higher for children with language disorders, and general developmental levels were low in the disordered groups. A total of 113 children were matched

Author contributions: D.K.O., P.N., S.G., J.A.R., J.G., D.X., and S.F.W. designed research; S.G., J.A.R., J.G., D.X., and U.Y. performed research; P.N., S.G., J.A.R., and D.X. contributed new reagents/analytic tools; D.K.O., P.N., S.G., J.A.R., and D.X. analyzed data; and D.K.O., P.N., and S.F.W. wrote the paper.

Conflict of interest statement: The recordings and hardware/software development were funded by Terrance and Judi Paul, owners of the previous for-profit company Infuture. Dissolution of the company was announced February 10, 2009, and it was reconstituted as the not-for-profit LENA Foundation. All assets of Infuture were given to the LENA Foundation. Before dissolution of the company, D.K.O., P.N., and S.F.W. had received consultation fees for their roles on the Scientific Advisory Board of Infuture. J.A.R., J.G., and D.X. are current employees of the LENA Foundation. S.G. and U.Y. are affiliates and previous employees of Infuture/LENA Foundation. None of the authors has or has had any ownership in Infuture or the LENA Foundation.

\*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence should be addressed. E-mail: koller@memphis.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1003882107/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1003882107/-DCSupplemental).



vocal characteristics would constitute a positive marker that might enhance precision of screening or diagnosis (31).

Most children with language-related disorders, even without autism, also show articulation and voice anomalies (36–39). Our work explores the possibility that the automated approach may be useful in discriminating typical from language-disordered development, as well as in discriminating autism from language delay. The power of the automated approach will presumably depend on its ability to predict changes with age in typically developing children.

## Results

Correlations between child age and SVI/SCU ratios for the 12 parameters showed clear evidence that the automated analysis predicts development: for the typically developing and language-delayed groups, five of the 12 correlations were greater than  $\pm 0.4$ , at a high level of statistical significance of  $P < 10^{-10}$  (Fig. 2*A* and *SI Appendix, Table S9 A–D*); all 12 correlations with age for the typically developing sample and seven of 12 for the language-delayed sample were statistically significant. The autistic sample, to the contrary, showed little evidence of development on the parameters; all their correlations of acoustic parameters with age were less than  $\pm 0.2$ , and only two were statistically reliable (*SI Appendix, Table S9 E and F*). The typically developing group showed reliable negative correlations ( $P < 10^{-4}$ ) on three parameters for which the other groups showed positive correlations, illustrating that certain vocal tendencies diminished with age in the typically developing group but did not diminish or increase with age in the others. Correlations of the 12 parameters with each other also revealed coherency within the four parameter groupings

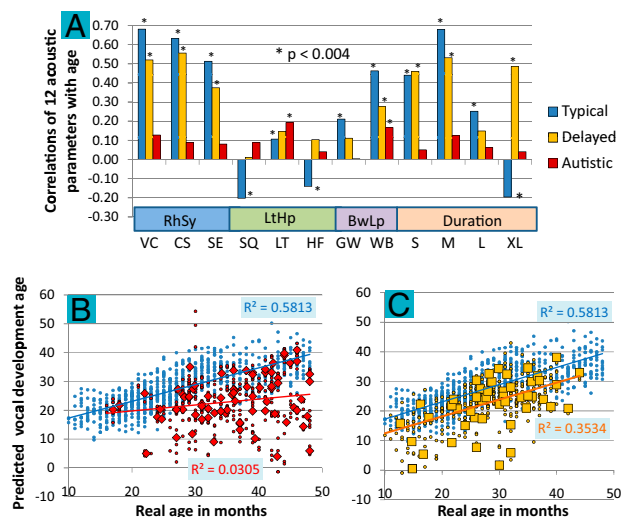
for all three child groups, but the autistic sample showed many more correlations not predicted by the four parameter groupings than the other child groups (*SI Appendix, Table S9 A–F*), a result suggesting children with autism organize acoustic infrastructure for vocalization differently from typically developing children or children with language delay.

Multiple linear regression (MLR) also illustrated development and group differentiation. The 12 acoustic parameter ratios (SVIs/SCUs) for each recording were regressed against age for the typically developing sample, yielding a normative model for development of vocalization. Coefficients of the model were used to calculate developmental ages for autistic and language-delayed recordings, displayed along with data for the typically developing recordings in Fig. 2*B* and *C*. The MLR model for the 802 recordings from the typically developing sample accounted for greater than 58% of variance in predicted age. MLR for the 106 typically developing children yielded yet higher prediction, accounting for more than 69% of variance, and suggesting that the automated analysis can provide a strong indicator of vocal development within this age range.

The correlation of predicted vocal developmental age with real age based on MLR was much higher for the typically developing and language-delayed samples (Fig. 2*C*) than for the autism sample (Fig. 2*B*), in which the predicted levels were also generally considerably below those of the typically developing sample. The acoustic parameters within the four theoretical groupings were also regressed against age (*SI Appendix, Fig. S11 A–D*), with results suggesting strong differentiation of the typically developing and autistic samples for three of the four groupings (RhSy, LtHp, and duration). The low correlations with age for the autism sample in these MLR analyses based on the typically developing model do not imply that the autism sample showed no change with regard to the acoustic parameters across age. On the contrary, an independent MLR model for the autism sample alone using all 12 acoustic parameters as predictors accounted for more than 16% of variance ( $P < 0.001$ ), indicating that the autism group did develop with age on the parameters. However the nature of the change with regard to the parameters across age was clearly different in the two groups, as indicated by the MLR results.

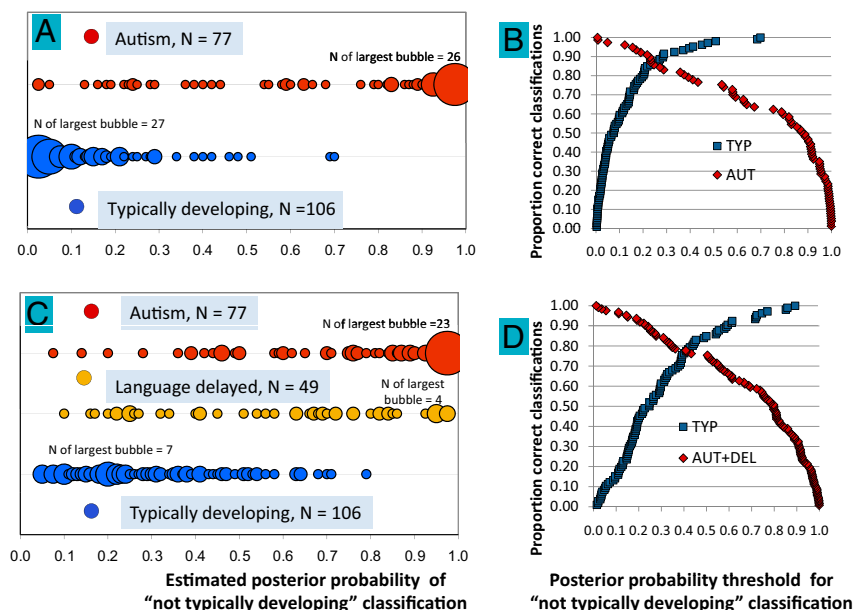
Linear discriminant analysis (LDA) and linear logistic regression were used to model classification of children into the three groups based on the acoustic parameters. Here the ratio scores (SVIs/SCUs) for each of the 12 parameters were first standardized by calculating z-scores for each recording in monthly age intervals. Leave-one-out cross-validation (LOOCV) was the main holdout method to verify generalization of results to recordings not in the training samples (*SI Appendix, “Statistical Analysis Summary”*). Six pairwise comparisons (*SI Appendix, Fig. S1*), each based on 1,500 randomized splits into training and testing (i.e., holdout) samples of varying sizes, confirmed that LOOCV results were consistently representative of midrange randomly selected holdout results, and thus provide a useful standard for the data reported in the subsequent paragraphs. LOOCV results for LDA and linear logistic regression were nearly identical (*SI Appendix, Table S1*); the remainder of the present report thus focuses on LDA only.

The results indicated strong correct identification of autism versus typical development ( $\chi^2$ ,  $P < 10^{-21}$ ), with sensitivity (i.e., hit rate) of 0.75 and specificity (i.e., correct rejection rate) of 0.98, based on a threshold (i.e., cutoff) posterior classification probability of 0.52 (Fig. 3*A*). At equal-error probability, sensitivity and specificity were 0.86 (crossover point, Fig. 3*B*). LDA also differentiated the typically developing sample from a combined autism and language-delay sample. At a cutoff of 0.56, specificity (i.e., correct identification of typically developing children) was 0.90. At the same cutoff, 80% of the autism sample and 62% of the language-delay sample were correctly classified as not being in the typically developing group (Fig. 3*C*). Sensitivity and specificity at equal error probability for classification of children as being or not



**Fig. 2.** Results of correlational analysis and MLR analysis. (*A*) Correlations of acoustic parameter SVI/SCU ratio scores with age across the 1,486 recordings. In 10 of 12 cases, both typically developing and language-delayed children showed higher absolute values of correlations with age than children with autism (*SI Appendix, Table S9 A–F*). All 12 correlations for the typically developing sample and seven of 12 for the language-delayed sample were statistically significant after Bonferroni correction ( $P < 0.004$ ). The autistic sample, conversely, showed little evidence of development on the parameters; all correlations of acoustic parameters with age were lower than  $\pm 0.2$ . (*B*) MLR for typically developing and autism samples. Blue dots represent real and predicted age (i.e., “predicted vocal development age”) for 802 recordings of typically developing children based on SVI/SCU ratios for the 12 acoustic parameters ( $r = 0.762$ ,  $R^2 = 0.581$ ). Red dots represent 351 recordings of children with autism, for which predicted vocal development ages ( $r = 0.175$ ,  $R^2 = 0.031$ ) were determined based on the typically developing MLR model. Each red diamond represents the mean predicted vocal development level across recordings for one of the 77 children with autism. (*C*) MLR for the typically developing (blue) and language-delayed samples (333 recordings; gold squares for 49 individual child averages;  $r = 0.594$ ,  $R^2 = 0.353$ ).





**Fig. 3.** LDA with LOOCV indicating differentiation of child groups. Estimated classification probabilities were based on the 12 acoustic parameters. Results are displayed in bubble plots, with each bubble sized in proportion to the number of children at each x axis location (the number of children represented by the largest bubble in each line is labeled). The plots indicate classification probabilities (*Left*) and proportion-correct classification (*Right*) for LDA in (A and B) children with autism (red) versus typically developing children (blue) and (C and D) a combined group of children with autism (red) or language delay (gold) versus typically developing children (blue).

being typically developing was 0.79 (Fig. 3D). Fig. 4A supplies LDA data on all six possible binary classificatory comparisons of the three child groups and illustrates strongest discrimination of the autism versus typically developing samples ( $P < 10^{-21}$ ). For the entire samples (left column), both typically developing and autism groups were also very reliably ( $P < 10^{-5}$ ) differentiated from the language-delayed group, although sensitivity/specificity was lower (0.73 and 0.70) than in the differentiation of the typically developing from autism groups or in differentiation of the typically developing from the combined language disordered groups.

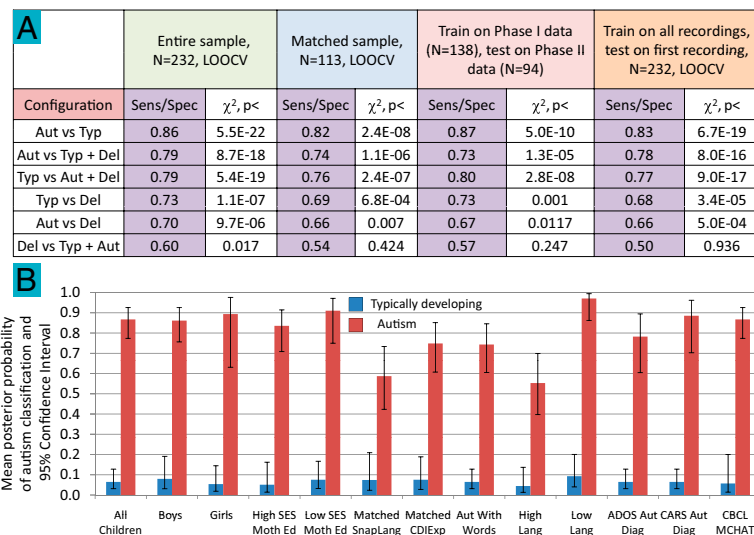
Fig. 44 also reveals only slightly lower levels of group differentiation for modeling with LOOCV on the matched sample of 113 children. Holdout models (i.e., not based on LOOCV) with training on phase I data and testing on phase II data revealed similar results. Models based on LOOCV training for all recordings were also used to determine sensitivity and specificity rates on only the first recording available for each child, again with similar results seen in Fig. 44. The first-recording analysis suggests the automated approach can discriminate groups quite reliably with a single all-day recording for each child.

Fig. 4B compares mean posterior probabilities of autism classification (LOOCV models) in subsamples for one configuration only: typically developing versus autism. Group differentiation was robust, applying to both genders and to higher and lower socioeconomic status (indicated by mother's reported education level). Further, differentiation was strong for subgroups matched on language development and expressive language measures, for the autism subgroup reported to be speaking in words, as well as for subgroups with high and low within-group language scores. Additional comparisons in Fig. 4B illustrate that subgroups diagnosed with autism through widely used tests documented in clinical reports supplied by parents (*SI Appendix, Table S3*) were differentiated sharply from the typically developing sample by the automated analysis. All comparisons of typically developing and autistic samples in Fig. 4B were highly statistically reliable ( $P < 10^{-11}$ ). Additional evaluations of posterior probabilities for individuals and subgroups of special interest in the autism and language-delay samples further illustrated the robustness of the automated analysis

in group differentiation (*SI Appendix*, “Results: Individual Children and Subgroups of Particular Interest in the Group Discrimination Analyses”; and *SI Appendix*, Figs. S14 and S15). Very low correlations of age with posterior probabilities ( $r = -0.07$  and  $r = 0.11$  for autism and typical samples, respectively) suggest group differentiation by this automated method was not importantly dependent on age within these samples (*SI Appendix*, Fig. S16).

Long-term success of automated vocal analysis may depend on the “transparency” of modeling, the extent to which results are interpretable through vocal development theory. The present work was based on such a theory (*SI Appendix*, “The 12 Acoustic Parameters” and “Research in Development of Vocal Acoustic Characteristics”), and thus provides a beginning for explaining why the model worked in, for example, predicting age. In pursuit of such explanation we further evaluated correlations among SVI/SCU ratios for the 12 acoustic parameters. To alleviate confounding effects of intercorrelations among the parameters, it was useful to project the data along orthogonal principal components (PCs). The PCA (*SI Appendix*, Table S10 and Fig. S12 A–C) yielded one dominant PC in the typically developing sample (analysis conducted on all 802 recordings without differentiation of the recordings within child), accounting for 40% of variance, more than double that of any other PC. The dominant PC for the typically developing group was highly correlated with all of the a priori RhSy (rhythm/syllabicity) parameters: (*i*) canonical transitions ( $r = 0.88$ ), reflecting for each recording the extent to which “syllables” (i.e., SVIs) had rapid formant transitions required in mature syllables, (*ii*) voicing (VC;  $r = 0.78$ ), indicating whether SVIs had periodic vocal energy required for mature syllables, and (*iii*) spectral entropy ( $r = 0.71$ ), indicating whether variations in vocal quality of SVIs were typical of variations occurring in mature speech. The dominant PC was also highly correlated with a parameter from the a priori duration grouping: (*iv*) medium duration ( $r = 0.91$ ), reflecting SVI durations in the mid-range for mature speech syllables.

When the four PCs of the typically developing sample were regressed against age at the child level, this dominant PC played the central role in age prediction ( $R^2 = 0.55$ ,  $P < 10^{-19}$ ). The  $\beta$ -coefficient (0.73) for the dominant PC was more than three



**Fig. 4.** LDA showing group discrimination for various configurations and subsamples. (A) LDA data on the six binary configurations of the three groups for the entire sample ( $N = 232$ ; LOOCV); subsamples matched on gender, mother's education level, and developmental level ( $n = 113$ ; LOOCV); a different holdout method in which, instead of LOOCV, training was conducted on phase I (2006–2008) data ( $n = 138$ ) and testing on phase II (2009) data ( $n = 94$ ); and a testing sample based for each child on the first recording only ( $N = 232$ ; LOOCV). (B) Bar graph comparisons for subsamples illustrating robustness of group differentiation in the autism versus typical development configuration only (based on LOOCV modeling). Means were calculated over logit-transformed posterior probabilities (PPs) of autism classification, then converted back to PPs. All comparisons showed robust group differentiation, including (from left to right) the entire sample ( $N = 232$ ), boys (typical,  $n = 48$ ; autism,  $n = 64$ ), girls (typical,  $n = 58$ ; autism,  $n = 13$ ), children of higher socioeconomic status (SES) as indicated by mother's education level  $\geq 6$  on an eight-point scale (typical,  $n = 42$ ; autism,  $n = 49$ ) and lower SES (typical,  $n = 64$ ; autism,  $n = 28$ ). To assess the possibility that "language level" may have played a critical role in automated group differentiation, we compared 35 child pairs matched for developmental age (typical development group, mean age of 22.6 mo; autism group, mean age of 22.7 mo) on the Snapshot, a language/communication measure (*SI Appendix*, "Participant Groups and Recording Procedures"), and 46 child pairs matched on the raw score from a single subscale of the CDI (40), namely, the expressive language subscale (typical development group, mean score of 21.6; autism group, mean score of 21.5), and we found robust group differentiation on PPs. Similar results were obtained for 48 children in the autism sample whose parents reported they were using spoken words meaningfully and for typical and autism samples split at medians into subgroups of high or low language (High Lang, Low Lang) level on the Snapshot developmental quotient. A subsample of 29 children with autism for whom diagnosis had been based on the Autism Diagnostic Observation Schedule (ADOS) (41) and another of 24 children with autism diagnosed with the Childhood Autism Rating Scale (CARS) (42) also showed robust group discrimination for PPs. Finally, in phase II (typical,  $n = 30$ ; autism,  $n = 77$ ) administration of both the Child Behavior Checklist (CBCL) (43) and the Modified Checklist for Autism in Toddlers (MCHAT) (44) had supported group assignment based on diagnoses, and group differentiation of PPs for these children using the automated system was unambiguous.

times that of any other PC, suggesting that the typically developing children's control of the infrastructure for syllabification provided the primary basis for modeling of vocal development age (*SI Appendix*, "Principal Components Analysis Indicating Empirical and Theoretical Organization of the Parameters").

PCA at the child level for the autism sample showed a considerably different structure of relations between acoustic parameters and PCs, and importantly, the first PC for the autism sample showed low correlations with the a priori RhSy acoustic parameters (*SI Appendix*, Table S10 and Fig. S12). When the four PCs for the autism sample were regressed against age, the PC with high correlations on the a priori RhSy parameters (PC2 in this case) yielded a  $\beta$  of 0.16, less than one fourth of that occurring for the typically developing sample. Prediction of age by this autism PC was not statistically significant. Thus, in the typically developing sample, the PC most associated with control of the RhSy parameters was highly predictive of age, but in the autism sample, no such predictiveness was found.

To quantify the role of RhSy parameters in group differentiation, the PCA model for the first four PCs in the typically developing sample was applied to child-level data from all three groups and then converted to standard scores based on the mean and SD for the typically developing group. Analysis of the mean standard scores (*SI Appendix*, Fig. S13) showed that, indeed, the PC associated with RhSy accounted for more than 3.5 times as much variance in group differentiation (adjusted  $R^2 = 0.21$ ) as any of the other PCs, suggesting that children's control of the infrastructure for syllabification may have played the predominant role in group differentiation.

## Discussion

The purpose of our work is fundamentally scientific: to develop tools for large-scale evaluation of vocal development and for research on foundations for language. The outcomes indicate that the time for basic research through automated vocal analysis of massive recording samples is already upon us. A typically developing child's age can be predicted at high levels (accounting for more than two thirds of variance) based on this first attempt at totally automated acoustic modeling with all-day recordings. The automated procedure has proven sufficiently transparent to offer suggestions on how the developmental age prediction worked—the primary factor appears to have been the child's command of the infrastructure for syllabification, a finding that should help to guide subsequent inquiries. From a practical standpoint, the present work illustrates the possibility of adding a convenient and fully objective evaluation of acoustic parameters to the battery of tests that is used with increasing success to identify children with language delay and autism. It appears that childhood control of the infrastructural features of syllabification played the central role in group differentiation, just as it played the central role in tracking development. The automated method showed higher accuracy in differentiating children with and without a language disorder than in differentiating the two language disorder groups (autism and language delay) from each other. Future work will be directed to evaluation of additional vocal factors that may differentiate subgroups of language disorder more effectively.

Based on the results reported here, there appears to be little reason for doubt that totally automated analysis of well selected acoustic features from naturalistic recordings can provide a mon-

itoring system for developmental patterns in vocalization as well as significant differentiation of children with and without disorders such as autism or language delay. We are optimistic that the procedure can be improved, as this is our first attempt at automated infrastructural modeling, with all parameters designed and implemented in advance of any analysis of recordings. To date there have still been no adjustments made in the theoretically proposed parameters. Additional modeling (e.g., hierarchical, age-specific, nonlinear) can be invoked, modifications can be explored in acoustic features, and larger samples, especially from the entire spectrum of ASDs and other language-related disorders, can help tune the procedures. The future of research on vocal development will profit from the combination of traditional approaches using laboratory-based exhaustive analysis of small samples of vocalization with the enormous power that is now clearly possible to command through automated analysis of naturalistic recordings.

- de Villiers JG, de Villiers PA (1974) Competence and performance in child language: Are children really competent to judge? *J Child Lang* 1:11–22.
- Slobin D (1973) Cognitive prerequisites for the development of grammar. *Studies in Child Language Development*, eds Ferguson CA, Slobin DI (Holt, Rinehart & Winston, New York), pp 175–208.
- Bloom L (1970) *Language Development* (MIT Press, Cambridge, MA).
- Brown R (1973) *A First Language* (Academic Press, London).
- Pinker S (1994) *The Language Instinct* (Harper Perennial, New York).
- Cutler A, Klein W, Levinson SC (2005) The cornerstones of twenty-first century psycholinguistics. *Twenty-First Century Psycholinguistics: Four Cornerstones*, ed Cutler A (Erlbaum, Mahwah, NJ), pp 1–20.
- Locke JL (1983) *Phonological Acquisition and Change* (Academic Press, New York).
- Oller DK (1980) The emergence of the sounds of speech in infancy. *Child Phonology, Vol 1: Production*, eds Yeni-Komshian G, Kavanagh J, Ferguson C (Academic Press, New York), pp 93–112.
- Stark RE, Rose SN, McLagen M (1975) Features of infant sounds: The first eight weeks of life. *J Child Lang* 2:205–221.
- Sheinkopf SJ, Mundy P, Oller DK, Steffens M (2000) Vocal atypicalities of preverbal autistic children. *J Autism Dev Disord* 30:345–354.
- Wetherby AM, et al. (2004) Early indicators of autism spectrum disorders in the second year of life. *J Autism Dev Disord* 34:473–493.
- Oller DK, Eilers RE (1988) The role of audition in infant babbling. *Child Dev* 59: 441–449.
- Eilers RE, Oller DK (1994) Infant vocalizations and the early diagnosis of severe hearing impairment. *J Pediatr* 124:199–203.
- Masataka N (2001) Why early linguistic milestones are delayed in children with Williams syndrome: Late onset of hand banging as a possible rate-limiting constraint on the emergence of canonical babbling. *Dev Sci* 4:158–164.
- Oller DK, Griebel U (2008) The origins of syllabification in human infancy and in human evolution. *Syllable Development: The Frame/Content Theory and Beyond*, eds Davis B, Zajdo K (Lawrence Erlbaum and Associates, Mahwah, NJ), pp 368–386.
- Vihman MM (1996) *Phonological Development: The Origins of Language in the Child* (Blackwell Publishers, Cambridge, MA).
- Stark RE, Ansel BM, Bond J (1988) Are prelinguistic abilities predictive of learning disability? A follow-up study. *Preschool Prevention of Reading Failure*, eds Masland RL, Masland M (York Press, Parkton, MD).
- Stoel-Gammon C (1989) Prespeech and early speech development of two late talkers. *First Lang* 9:207–223.
- Oller DK (2000) *The Emergence of the Speech Capacity* (Lawrence Erlbaum Associates, Mahwah, NJ).
- Kanner L (1943) Autistic disturbances of affective contact. *Nerv Child* 2:217–250.
- Asperger H (1991) Autistic “psychopathy” in childhood. *Autism and Asperger Syndrome*, ed Frith U (Cambridge University Press, Cambridge, UK), pp 37–90.
- Paul R, Augustyn A, Klin A, Volkmar FR (2005) Perception and production of prosody by speakers with autism spectrum disorders. *J Autism Dev Disord* 35:205–220.
- Pronovost W, Wakstein MP, Wakstein DJ (1966) A longitudinal study of the speech behavior and language comprehension of fourteen children diagnosed atypical or autistic. *Except Child* 33:19–26.
- McCann J, Peppé S (2003) Prosody in autism spectrum disorders: a critical review. *Int J Lang Commun Disord* 38:325–350.

## Materials and Methods

The following details of methods are provided in *SI Appendix, SI Materials and Methods*: (i) statistical analysis summary, (ii) recording device, (iii) participant groups and recording and assessment procedures, (iv) automated analysis algorithms, (v) the 12 acoustic parameters, and (vi) reliability of the automated analysis. The statistical analysis summary provides details regarding methods of prediction of development using MLR, as well as differentiation of child groups using LDA and logistic regression. An empirical illustration of reasons for using LOOCV in LDA and logistic regression analyses is also given. The recording procedures used in the present work as well as recruitment procedures for children in the three groups are explained along with selection criteria. Additional data are supplied on language and social characteristics of the participating groups. Automated analysis algorithms are described along with reliability data. All procedures were approved by the Essex Institutional Review Board.

**ACKNOWLEDGMENTS.** Research by D.K.O. for this paper was funded by an endowment from the Plough Foundation, which supports his Chair of Excellence at The University of Memphis.

- Peppé S, McCann J, Gibbon F, O'Hare A, Rutherford M (2007) Receptive and expressive prosodic ability in children with high-functioning autism. *J Speech Lang Hear Res* 50:1015–1028.
- Shriberg LD, et al. (2001) Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *J Speech Lang Hear Res* 44: 1097–1115.
- Association AP (2000) *Diagnostic and Statistical Manual of Mental Disorders, DSM-IV-TR* (American Psychiatric Association, Arlington, VA).
- Baron-Cohen S, Allen J, Gillberg C (1992) Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *Br J Psychiatry* 161:839–843.
- Baron-Cohen S (2000) Theory of mind and autism: a fifteen year review. *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*, eds Baron-Cohen S, Tager-Flusberg H, Cohen DJ (Oxford University Press, Oxford), pp 3–20.
- Loveland KA, Landry SH (1986) Joint attention and language in autism and developmental language delay. *J Autism Dev Disord* 16:335–349.
- Mundy P, Kasari C, Sigman M (1992) Nonverbal communication, affective sharing, and intersubjectivity. *Infant Behav Dev* 15:377–381.
- Mundy P, Sigman M, Kasari C (1990) A longitudinal study of joint attention and language development in autistic children. *J Autism Dev Disord* 20:115–128.
- Rutter M (1978) Diagnosis and definition of childhood autism. *J Autism Child Schizophr* 8:139–161.
- Tager-Flusberg H (1989) A psycholinguistic perspective on language development in the autistic child. *Autism: Nature, Diagnosis, and Treatment*, ed Dawson G (Guilford, New York), pp 92–115.
- Tager-Flusberg H (2000) Language and understanding minds: connections in autism. *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*, eds Baron-Cohen S, Tager-Flusberg H, Cohen DJ (Oxford University Press, Oxford), pp 124–149.
- Tyler AA, Sandoval KT (1994) Preschoolers With Phonological and Language Disorders: Treating Different Linguistic Domains. *Lang Speech Hear Serv Schools* 25:215–234.
- Conti-Ramsden G, Botting N (1999) Classification of children with specific language impairment: Longitudinal considerations. *J Speech Lang Hear Res* 42:1195–1204.
- Shriberg LD, Aram DM, Kwiatkowski J (1997) Developmental apraxia of speech: I. Descriptive and theoretical perspectives. *J Speech Lang Hear Res* 40:273–285.
- Marshall CR, Harcourt-Brown S, Ramus F, van der Lely HK (2009) The link between prosody and language skills in children with specific language impairment (SLI) and/or dyslexia. *Int J Lang Commun Disord* 44:466–488.
- Ireton H, Glascoe FP (1995) Assessing children's development using parents' reports. The Child Development Inventory. *Clin Pediatr (Phila)* 34:248–255.
- Lord C, Rutter M, DiLavore PC, Risi S (2002) *Autism Diagnostic Observation Schedule* (Western Psychological Services, Los Angeles).
- Schopler E, Reichler RJ, DeVellis RF, Daly K (1980) Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *J Autism Dev Disord* 10: 91–103.
- Achenbach TM (1991) *Integrative Guide to the 1991 CBCL/4-18, YSR, and TRF Profiles* (University of Vermont, Burlington, VT).
- Robins DL, Fein D, Barton ML, Green JA (2001) The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *J Autism Dev Disord* 31:131–144.