Machine Learning Identifies Infant Speech Preference as an Important Predictor for Early Autism Risk Detection

Yuhan Jiang
School of Foreign Studies,
Tongji University, Shanghai 200092, China
Qiang Fu
Center for Speech and Language Processing,
Tongji University, Shanghai 200092, China

The University of Osaka - Tongji University
Joint International and Interdisciplinary Forum:
Toward the Integration of Society and Knowledge
Official Forum Proceedings

Abstract

While conventional ASD risk assessments rely on later-emerging behavioral features, early vocal preferences offer infant-available predictive markers. Therefore, multimodal measures integrating behavioral features with vocal preference may enhance autistic risk prediction accuracy. This study developed machine learning models to identify high autistic risk infants using multimodal measures, including gaze duration of sound contrasts (speech vs. non-vocal; speech vs. monkey calls) and learning/motor scores. Analysis of 9/12 month infant data (56 high-risk, 82 low-risk) showed SVM performed best after optimization (71.43% accuracy, AUC=0.46). Average model feature importance analysis identified that key predictors are 9-month human-over-monkey preference, 9-month speech-over-nonspeech preference, and 12-month receptive language percentiles. Findings demonstrate that gaze-based speech preference assessment addresses critical limitations of traditional behavioral markers, and when integrated with predictive modeling, opens new pathways for significant advances in early autism risk detection.

Keywords: Machine learning, Early autistic risk identification, Multimodal behaviors, Speech preference

Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by persistent deficits in verbal/non-verbal communication, impaired social reciprocity, and restricted/repetitive behavioral patterns (American Psychiatric Association, 2013). Symptom severity and functional impact demonstrate pronounced individual heterogeneity across all racial, ethnic, and socioeconomic groups (Alsharif et al., 2024). Crucially, attenuated behavioral manifestations in specific subpopulations (notably females) frequently evade timely clinical recognition compared to more overt presentations (e.g., in males), resulting in significant diagnostic delays (Solomon et al., 2012). This discrepancy occurs despite equivalent core social-communication impairments, thereby creating intervention access disparities that compound functional challenges throughout the lifespan. Consequently, early identification remains a clinical imperative, as evidence-based interventions initiated during critical developmental windows substantially improve long-term adaptive outcomes. Beyond autistic children themselves, their infant siblings, who face substantially elevated autism risk (≈19% diagnosis rate by age 3 versus ≈1.5% in the general population) (Ozonoff et al., 2011), likewise warrant careful monitoring from infancy.

Current approaches to early autism diagnosis primarily rely on identifying behavioral markers emerging in toddlerhood (Hyman et al., 2020). However, certain infant signals, such as speech sound bias, may offer earlier indicators of autism risk. This is critical because infants must distinguish meaningful speech from background noise to acquire communication skills (Liberman & Whalen, 2000; Pinker & Jackendoff, 2005; Sorcinelli et al., 2019) and support their language and social development (Vouloumanos & Werker, 2007; Werker & Curtin, 2005). Typically developing (TD) newborns exhibit a natural bias towards speech over many other sounds, a preference believed to underpin essential learning processes (Kuhl, 1988; Vouloumanos & Werker, 2007). Crucially, speech processing ability in infancy consistently predicts later language and social skills (Cristia et al., 2014), including joint attention (Roberts et al., 2013). While autistic children often lack typical speech preferences, the degree of speech bias they do exhibit nevertheless predicts receptive language ability (Paul et al., 2007) and sensitivity to syllable changes (Kuhl et al., 2005). Furthermore, infant preferences for speech over non-biological non-speech sounds show differential predictive value for later developmental outcomes among siblings of autistic children (Curtin & Vouloumanos, 2013; Vouloumanos & Curtin, 2014).

Eye tracking, a non-invasive measurement technique, offers enhanced precision in detecting atypical visual patterns in infants later diagnosed with autism (Papagiannopoulou et al., 2014). Substantial evidence links autism to aberrant eye-tracking behaviors (Senju & Johnson, 2009).

In infant siblings of autistic children, measuring gaze duration toward different auditory stimuli complements traditional behavioral assessments in predicting subsequent linguistic development, social communication skills, and autism-related behaviors (Sorcinelli et al., 2019). Furthermore, eye tracking can be effectively combined with measures of speech bias, as gaze patterns directly reveal subtle variations in young children’s auditory attention.
Building on the potential of combining speech preference and gaze behaviors for early detection, machine learning offers a promising approach to further enhance autism risk identification. By providing an unbiased and reproducible secondary assessment (Sasson & Elison, 2012; Yaneva et al., 2020), machine learning can integrate these multimodal signals with other observational behaviors. Therefore, this study investigates whether machine learning algorithms and gaze patterns reflecting speech preference can effectively facilitate traditional behavioral features in distinguishing infant siblings of autistic children from siblings of TD children. This distinction is based on combined measures of: (1) established behavioral and language assessments (Autism Observation Scale for Infants, AOSI; MacArthur-Bates Communicative Development Inventories, CDI; Mullen Scales of Early Learning, MSEL), and (2) gaze patterns reflecting auditory preferences during exposure to speech, non-biological non-speech sounds, and monkey calls.

Our predictor framework capitalizes on complementary developmental windows to capture emergent risk trajectories. At 9 months, auditory processing features, including gazing responses to speech and non-speech stimuli, reveal pre-symptomatic tuning atypicality, while concurrent behavioral observations contextualize risk before language onset. By 12 months, standardized language assessments quantify emerging communicative deficits alongside behavioral features that track consolidating autism characteristics as diagnostic profiles. This staged multimodal design aligns with established neurodevelopmental cascades, wherein early auditory anomalies precede observable symptoms, which subsequently manifest as measurable social-communication and language delays (Sorcinelli et al., 2019). Given the study’s feature dimensionality and sample size, we posit that 1) models robust to limited data will optimally leverage these temporally stratified inputs; 2) 9-month speech processing metrics, particularly auditory preference for speech, will predict autistic risk comparable to conventional linguistic and behavioral markers.

Methods

Dataset

The dataset was obtained from a public repository (Sorcinelli et al., 2018) containing eye-tracking measures from 9-month-old infants during exposure to speech, non-biological non-speech sounds, and monkey calls. Participants included siblings of autistic children (SIBS-A;

27 male, 29 female) and siblings of TD children (SIBS-TD; 38 male, 44 female). The participants without group or gender information were excluded. Missing values in gaze duration measures were addressed through median imputation.

Auditory stimuli comprised monkey calls, monosyllabic nonsense speech words (‘lif’, ‘neem’; 6 tokens each) spoken by a female native English speaker (pitch: 197-350 Hz; duration: 525-1155 ms), and acoustically matched non-speech analogues (Vouloumanos et al., 2001; Vouloumanos & Werker, 2004). These analogues were time-varying sinusoidal waves replicating the fundamental frequency and first three formants, preserving the duration, pitch contour, amplitude envelope, relative formant amplitude, and relative intensity of the speech tokens. Crucially, non-speech analogues differed in lacking natural voice quality and biological realism and utilizing four independent sound sources versus the single vocal tract source of speech. Infants were tested using an infant-controlled preferential looking procedure in a sound-attenuated room (sound level: 60 ± 5 dB). Seated on a parent’s lap 89 cm from a monitor, infants initiated trials by fixating a central stimulus (Cooper & Aslin, 1990; Vouloumanos & Werker, 2004). Looking time to the screen, provided by the dataset publisher, was coded offline frame-by-frame (30 fps) by observers being blind to trial type; reliability was high (ICC = .98, $p < .001$, based on 20% double-coding). Supplementary measures included the AOSI (9/12 months) (Bryson et al., 2008), MB-CDI (12 months) (Fenson et al., 1993), and MSEL (12 months) (Mullen, 1995).

Machine Learning Models

A predictive modeling framework was developed to classify autism risk groups (SIBS-TD versus SIBS-A) using behavioral features measured during infancy. The analytical approach incorporated systematic data preprocessing and multiple classification algorithms. Missing values were addressed through k-nearest neighbors imputation (k=5), followed by robust scaling to minimize outlier effects. Class imbalance was systematically managed through differential weighting and synthetic minority oversampling during cross-validation. Five classification algorithms were implemented: 1) XGBoost with asymmetric error weighting (Chen & Guestrin, 2016), 2) LightGBM with gradient-based sampling (Ke et al., 2017), 3) Random Forest with balanced class weighting (Breiman, 2001), 4) Support Vector Machine with radial basis kernel (Chang & Lin, 2011), and 5) Multilayer Perceptron with early stopping regularization (Rumelhart et al., 1986).

Model development employed stratified 3-fold cross-validation with hyperparameter optimization through grid search. Balanced accuracy, defined as the mean of sensitivity and specificity, served as the primary model selection criterion. The support vector machine emerged as the optimal classifier based on test set AUC performance. The three top-performing classifiers were subsequently integrated into a stacked ensemble using logistic regression as a meta-classifier, validated through 5-fold cross-validation.

Performance evaluation encompassed accuracy, AUC, and F1 scores. Model interpretation included feature importance assessment through Gini impurity decomposition for compatible tree-based algorithms (XGBoost, LightGBM, and Random Forest). Aggregate importance scores were computed by averaging across these models, providing insights into key behavioral predictors. Classification thresholds were optimized via F1-score maximization for the optimal classifier. Diagnostic visualizations included receiver operating characteristic curves, precision-recall plots, and confusion matrices. The computational implementation utilized Python with scikit-learn (v1.0), XGBoost (v1.5), and LightGBM (v3.3) libraries.

Results

The machine learning models were evaluated on an independent test set representing 15% of the overall cohort (n=138). Among the evaluated approaches, the Support Vector Machine (SVM) demonstrated superior classification accuracy at 71.43%, substantially outperforming alternative methods (Table 1).

Table 1. Summary of models' outputs

	Accuracy	AUC	F1
XGBoost	38.10%	0.38	0.32
LightGBM	38.10%	0.32	0.32
Random Forest	47.62%	0.40	0.42
SVM	71.43%	0.46	0.63
Multilayer Perceptron	57.14%	0.64	0.40
Ensemble	42.86%	0.42	0.40

The SVM maintained robust performance across both risk groups, correctly classifying the majority of cases despite the complexity of predicting developmental outcomes from infant multi-modal features. The SVM also achieved the highest F1-score (0.63), indicating balanced precision and recall in identifying both risk groups (Figure 1).

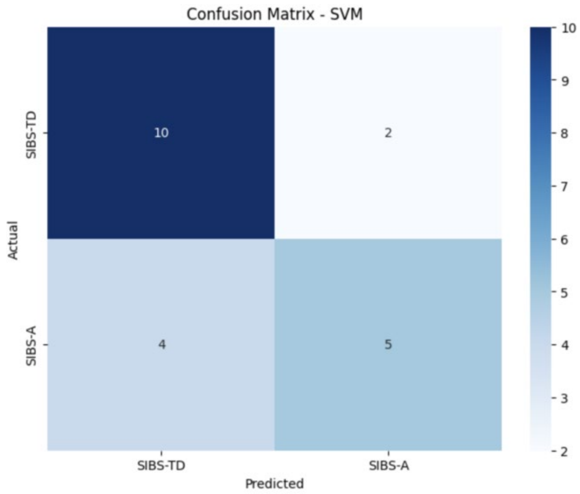

Figure 1: Confusion matrix of the SVM model.

Notably, the Multilayer Perceptron model showed the highest AUC (0.639), suggesting better risk probabilities. However, this apparent advantage did not translate to practical clinical utility, as evidenced by its lower accuracy (57.14%) and substantially reduced F1-score (0.40). This performance pattern indicates potential calibration issues where probability estimates didn't consistently correspond to correct classifications. The ensemble approach, while theoretically promising, failed to outperform the best individual classifier in this application.

Figure 2 presents the average feature importance derived from three tree-based models (XGBoost, LightGBM, and RandomForest) that support intrinsic importance metrics. The hierarchical ranking reveals the preference for Human speech over monkey calls at 9-month-old as the most influential predictor, demonstrating substantially higher importance than other features. This is closely followed by the preference for Human speech over non-speech calls at 9-month-old and the receptive language percentile of MSEL at 12-month-old, indicating that speech preference measures and receptive language abilities constitute critical predictive signals for autism risk classification.

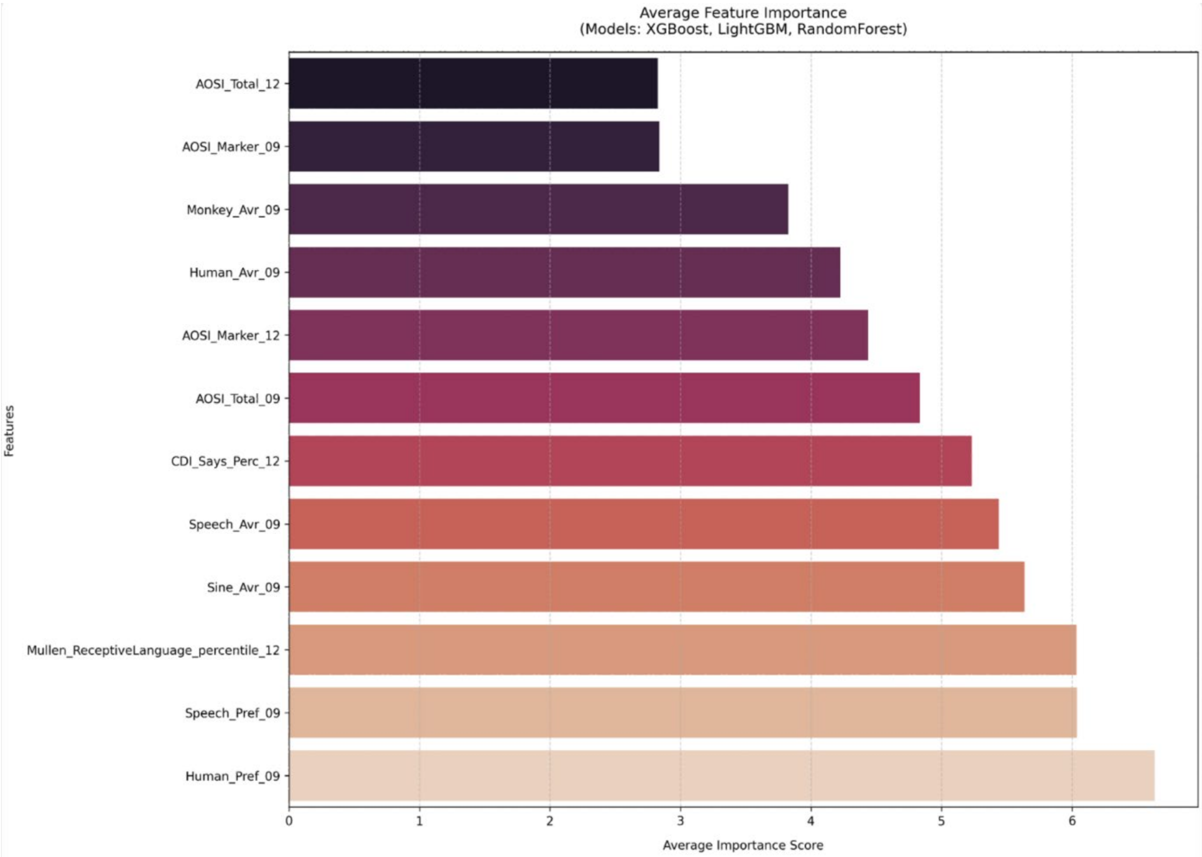

Figure 2: Average feature importance of the three tree-based models ('09/12' for the age in months; 'Avr' for average looking time during speech trials; 'Perc' for Percentile; 'total' for total scores of 18 items in AOSI, indicating severity; 'marker' for total number of AOSI items scored between 1 and 3, indicating the breadth of symptom expression; 'Pref' for preference, i.e., looking time difference between different trials).

Mid-range importance features include some auditory processing metrics (average looking time during Sine and speech trials at 9 months old), Expressive language ability (CDI Says percentile at 12 months old), and autism observational scores (severity of symptoms at 9 months old and breadth of symptoms at 12 months old). Notably, other early auditory responses (average looking time during Human and Monkey trials at 9 months old) showed moderate predictive value. The symptom breadth at 9 months and symptom severity at 12 months showed lower importance.

Discussion

This study investigated the utility of machine learning algorithms in differentiating infant siblings of autistic children from siblings of typically developing children using integrated behavioral assessments and gaze-based auditory preference measures. Five classification models were evaluated: XGBoost, LightGBM, Random Forest, Support Vector Machine, and

Multilayer Perceptron. The Support Vector Machine achieved superior classification performance and the highest F1-score. Feature importance analysis across tree-based models identified the preference for human speech over monkey calls at 9 months as the most influential predictor. This was closely followed by the preference for speech over non-speech sounds at 9 months and receptive language abilities at 12 months. Moderately important predictors included expressive language skills at 12 months, some auditory processing metrics, autism observational total score at 9 months, and autism observational marker at 12 months. Autism observational marker at 9 months and autism observational total score at 12 months demonstrated relatively lower predictive importance.

Performance of Machine Learning Models

Among the evaluated models, the SVM model demonstrated superior performance across key clinical metrics, achieving the highest accuracy and F1-score in differentiating SIBS-A and SIBS-TD infants. The performance divergence between SVM and other models stems from core algorithmic differences that become particularly consequential in autism risk prediction. While the Multilayer Perceptron showed marginally better ranking capability, its substantially lower accuracy and F1-score reveal fundamental limitations in converting risk stratification into actionable classifications. This calibration gap reflects neural networks' requirement for larger samples to translate probabilistic sophistication into reliable decisions (Alwosheel et al., 2018), whereas SVM's structural risk minimization provides stronger generalization for small-sample, low-dimensional data, where its training complexity remains manageable due to limited observations (Yu et al., 2003). These differences are exacerbated in simple binary autism risk prediction, where feature scarcity and limited samples amplify SVM's suitability. This performance pattern aligns with established machine learning principles, where complex models often underperform some simpler alternatives in data-constrained scenarios like our infant sibling cohort (Alarifi et al., 2023).

In infant autism sibling screening, the narrow window for effective interventions (Vivanti et al., 2019) highlights the need to prioritize reliable classification over marginal improvements in ranking metrics. False negatives carry irreversible consequences, making sensitivity-optimized accuracy a non-negotiable priority. Simultaneously, F1-score balances this sensitivity against precision, reducing false positives that strain clinical resources and cause familial distress. While AUC measures probabilistic ranking quality, its clinical limitations are mitigated by clinician-led secondary assessments. SVM serves as a potentially operational decision-support tool, identifying infants warranting comprehensive evaluation while acknowledging its probabilistic nature. This framework stresses actionable performance (accuracy/F1) over theoretical ranking sophistication (AUC), as timely identification with managed false positives aligns with the ethical imperative of preventing missed interventions.

Speech preference as a predictive feature of autistic risk

Feature importance analysis confirmed the importance of speech processing, identifying 9-month auditory speech preference over monkey calls as the strongest predictor, followed by 9-month speech-nonspeech differentiation and 12-month receptive language. These results suggest that speech preference can be a reliable predictor for early autistic risk prediction, like other linguistic or behavioral symptoms. Preferences for speech early in infancy may play an important role in setting typical linguistic and social communicative development, particularly for infants with autistic siblings, who are themselves at higher risk for later being diagnosed with autism, developing language or other developmental delays, or showing subclinical features of autistic characteristics of the broader autism phenotype (Ozonoff et al., 2014).

The moderately weighted predictors, including 12-month expressive language ability, autism behavioral scores, and absolute looking time to auditory stimuli at 9 months, demonstrate relatively limited standalone value in predicting autistic risk. While isolated gaze duration measures provide some utility by indexing undifferentiated attention to auditory stimuli, they lack diagnostic specificity for autistic risk prediction, primarily due to their inability to capture infants' discriminative processing of socially meaningful sounds. In contrast, speech preference, measured through preferential attention to human speech versus nonspeech sounds or monkey calls, provides superior predictive power because it directly indexes early consciousness of the social auditory system (Vouloumanos & Curtin, 2014). When infants consistently favor speech over nonspeech/monkey calls, it reflects specialization for useful vocalizations (Vouloumanos & Werker, 2004), which in turn facilitates richer language exposure and responsive caregiving interactions (Warlaumont et al., 2014). Thus, while conventional measures offer reliable insights, speech-over-nonspeech and human-over-monkey preference provide new evidence for early autistic risk.

The symptom breadth at 9 months and symptom severity at 12 months showed lower importance in our predictive model. This finding should not be interpreted as diminishing the established clinical value of these behavioral measures, which remain essential diagnostic indicators across numerous studies (Hosozawa et al., 2020; Hyman et al., 2020). Rather, their reduced weighting reflects inherent developmental constraints: at 9 months, infants' limited behavioral repertoire constrains comprehensive symptom breadth assessment, while at 12 months, symptom severity measures capture phenotypes still undergoing dynamic consolidation. This is precisely where auditory preference measures offer critical complementary value. Speech-nonspeech and human-monkey contrast paradigms objectively quantify subtle attentional biases through gaze patterns toward socially meaningful sounds (Sorcinelli et al., 2019), circumventing reliance on overt infant behaviors that may be transient or situation-dependent. By detecting early neurodevelopmental signals months before robust behavioral symptoms stabilize, speech preference provides a unique window into emerging social communication tendency. Thus, while behavioral measures maintain their diagnostic significance, speech preference metrics enhance predictive sensitivity during this developmental period by capturing foundational responses that precede observable symptoms.

Limitation

This study's identification of machine learning's potential and speech preference as an early predictor advances autism risk detection, yet has several limitations. First, findings derived exclusively from the small and low-dimensional sample limit generalizability to screening contexts. Second, our cross-sectional assessment captures a single developmental snapshot, which cannot characterize dynamic risk trajectories across infancy. Third, though multimodal feature integration demonstrated superior predictive validity over single-modality approaches, the probabilistic risk estimates remain adjunctive to, not replacements for, comprehensive clinical evaluation. Future validation should increase the sample size, include more features, and develop interpretable interfaces that translate feature importance into clinical workflows.

Conclusion

This study demonstrates that infants' preference for human speech, measured through their gaze patterns when hearing speech versus other sounds, serves as a powerful early predictor of autism risk. Machine learning analysis confirmed that this speech preference outperformed traditional developmental measures in identifying infants with higher autistic risk. While later language skills and behavioral observations provided valuable insights, the simple act of observing whether infants consistently favored listening to people over other sounds proved most revealing. This discovery highlights how fundamental auditory social attention is to early development, and how its measurement creates new opportunities to identify infants needing support, long before clear behavioral symptoms emerge. Combining machine learning with speech preference measurements opens new pathways for significant advances in early autism risk detection.

Data statement

Acknowledgements

We would like to express our sincere gratitude to the dataset developers and the children and families who participated in the development of the dataset.

References

Alarifi, H., Aldhalaan, H., Hadjikhani, N., Johnels, J. Å., Alarifi, J., Ascenso, G., & Alabdulaziz, R. (2023). Machine learning for distinguishing saudi children with and without autism via eye-tracking data. *Child and Adolescent Psychiatry and Mental Health*, *17*(1), 112. https://doi.org/10.1186/s13034-023-00662-3

Alsharif, N., Al-Adhaileh, M. H., Al-Yaari, M., Farhah, N., & Khan, Z. I. (2024). Utilizing deep learning models in an intelligent eye-tracking system for autism spectrum disorder diagnosis. *Frontiers in Medicine*, *11*. https://doi.org/10.3389/fmed.2024.1436646

Alwosheel, A., van Cranenburgh, S., & Chorus, C. G. (2018). Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling*, *28*, 167–182. https://doi.org/10.1016/j.jocm.2018.07.002

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). American psychiatric association. https://www.psychiatry.org/psychiatrists/practice/dsm

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bryson, S. E., Zwaigenbaum, L., McDermott, C., Rombough, V., & Brian, J. (2008). The Autism Observation Scale for Infants: Scale Development and Reliability Data. *Journal of Autism and Developmental Disorders*, *38*(4), 731–738. https://doi.org/10.1007/s10803-007-0440-y

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, *2*(3), 27:1-27:27. https://doi.org/10.1145/1961189.1961199

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, 785–794. https://doi.org/10.1145/2939672.2939785

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584–1595.

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting Individual Variation in Language From Infant Speech Perception Measures. *Child Development*, *85*(4), 1330–1345. https://doi.org/10.1111/cdev.12193

Curtin, S., & Vouloumanos, A. (2013). Speech preference is associated with autistic-like behavior in 18-months-olds at risk for Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *43*(9), 2114–2120. https://doi.org/10.1007/s10803-013-1759-1

Fenson, L., Dale, P., Reznick, J., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. Singular Publishing Group.

Hosozawa, M., Sacker, A., Mandy, W., Midouhas, E., Flouri, E., & Cable, N. (2020). Determinants of an autism spectrum disorder diagnosis in childhood and adolescence: Evidence from the UK Millennium Cohort Study. *Autism*, *24*(6), 1557–1565. https://doi.org/10.1177/1362361320913671

Hyman, S. L., Levy, S. E., Myers, S. M., COUNCIL ON CHILDREN WITH DISABILITIES, S. O. D. A. B. P., Kuo, D. Z., Apkon, S., Davidson, L. F., Ellerbeck, K. A., Foster, J. E. A., Noritz, G. H., Leppert, M. O., Saunders, B. S., Stille, C., Yin, L., Weitzman, C. C., Childers, D. O., Jr, Levine, J. M., Peralta-Carcelen, A. M., Poon, J. K., … Bridgemohan, C. (2020). Identification, Evaluation, and Management of Children With Autism Spectrum Disorder. *Pediatrics*, *145*(1), e20193447. https://doi.org/10.1542/peds.2019-3447

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems*, *30*. https://proceedings.neurips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

Kuhl, P. K. (1988). Auditory perception and the evolution of speech. *Human Evolution*, *3*(1), 19–43. https://doi.org/10.1007/BF02436589

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science*, *8*(1), F1–F12. https://doi.org/10.1111/j.1467-7687.2004.00384.x

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, *4*(5), 187–196. https://doi.org/10.1016/S1364-6613(00)01471-6

Mullen, E. (1995). Mullen scales of early learning. In *Circle Pines*. American Guidance Service. https://www.sciencedirect.com/topics/psychology/mullen-scales-of-early-learning

Ozonoff, S., Young, G. S., Belding, A., Hill, M., Hill, A., Hutman, T., Johnson, S., Miller, M., Rogers, S. J., Schwichtenberg, A. J., Steinfeld, M., & Iosif, A.-M. (2014). The Broader Autism Phenotype in Infancy: When Does It Emerge? *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(4), 398-407.e2. https://doi.org/10.1016/j.jaac.2013.12.020

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Bryson, S., Carver, L. J., Constantino, J. N., Dobkins, K., Hutman, T., Iverson, J. M., Landa, R., Rogers, S. J., Sigman, M., & Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: A Baby Siblings Research Consortium study. *Pediatrics*, *128*(3), e488-495. https://doi.org/10.1542/peds.2010-2825

Papagiannopoulou, E. A., Chitty, K. M., Hermens, D. F., Hickie, I. B., & Lagopoulos, J. (2014). A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social Neuroscience*, *9*(6), 610–632. https://doi.org/10.1080/17470919.2014.934966

Paul, R., Chawarska, K., Fowler, C., Cicchetti, D., & Volkmar, F. (2007). “Listen My Children and You Shall Hear”: Auditory Preferences in Toddlers With Autism Spectrum Disorders. *Journal of Speech, Language, and Hearing Research*, *50*(5), 1350–1364. https://doi.org/10.1044/1092-4388(2007/094)

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What’s special about it? *Cognition*, *95*(2), 201–236. https://doi.org/10.1016/j.cognition.2004.08.004

Roberts, S., Fyfield, R., Baibazarova, E., Van Goozen, S., Culling, J. F., & Hay, D. F. (2013). Parental Speech at 6Months Predicts Joint Attention at 12Months. *Infancy*, *18*, E1–E15. https://doi.org/10.1111/infa.12018

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Sasson, N. J., & Elison, J. T. (2012). Eye Tracking Young Children with Autism. *Journal of Visualized Experiments (JoVE)*, (61), e3675. https://doi.org/10.3791/3675

Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127–134. https://doi.org/10.1016/j.tics.2008.11.009

Solomon, M., Miller, M., Taylor, S. L., Hinshaw, S. P., & Carter, C. S. (2012). Autism Symptoms and Internalizing Psychopathology in Girls and Boys with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *42*(1), 48–59. https://doi.org/10.1007/s10803-011-1215-z

Sorcinelli, A., Ference, J., Curtin, S., & Vouloumanos, A. (2019). Preference for speech in infancy differentially predicts language skills and autism-like behaviors. *Journal of Experimental Child Psychology*, *178*, 295–316. https://doi.org/10.1016/j.jecp.2018.09.011

Sorcinelli, A., Vouloumanos, A., Ference, J., & Curtin, S. (2018). *Data for: Preference for speech in infancy differentially predicts language skills and autism-like behaviors* (Version 1) [Dataset]. Mendeley Data. https://doi.org/10.17632/npw24yrjf6.1

Vivanti, G., Dissanayake, C., Duncan, E., Feary, J., Capes, K., Upson, S., Bent, C. A., Rogers, S. J., & Hudry, K. (2019). Outcomes of children receiving Group-Early Start Denver Model in an inclusive versus autism-specific setting: A pilot randomized controlled trial. *Autism*, *23*(5), 1165–1175. https://doi.org/10.1177/1362361318801341

Vouloumanos, A., & Curtin, S. (2014). Foundational Tuning: How Infants' Attention to Speech Predicts Language Development. *Cognitive Science*, *38*(8), 1675–1686. https://doi.org/10.1111/cogs.12128

Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience*, *13*(7), 994–1005. https://doi.org/10.1162/089892901753165890

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science*, *7*(3), 270–276. https://doi.org/10.1111/j.1467-7687.2004.00345.x

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159–164. https://doi.org/10.1111/j.1467-7687.2007.00549.x

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A Social Feedback Loop for Speech Development and Its Reduction in Autism. *Psychological Science*, *25*(7), 1314–1324. https://doi.org/10.1177/0956797614531023

Werker, J. F., & Curtin, S. (2005). PRIMIR: A Developmental Framework of Infant Speech Processing. *Language Learning and Development*, *1*(2), 197–234. (world). https://www.tandfonline.com/doi/abs/10.1080/15475441.2005.9684216

Yaneva, V., Ha, L. A., Eraslan, S., Yesilada, Y., & Mitkov, R. (2020). Detecting High-Functioning Autism in Adults Using Eye Tracking and Machine Learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28*(6), 1254–1261. https://doi.org/10.1109/TNSRE.2020.2991675

Yu, H., Yang, J., & Han, J. (2003). Classifying Large Data Sets Using SVMs with Hierarchical Clusters. *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 306–315.