



Infants' neural speech discrimination predicts individual differences in grammar ability at 6 years of age and their risk of developing speech-language disorders

T. Christina Zhao^{a,b,*}, Olivia Boorom^c, Patricia K. Kuhl^{a,b}, Reyna Gordon^{d,e,f}

^a Institute for Learning & Brain Sciences, University of Washington, Seattle, WA, USA

^b Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, USA

^c Department of Hearing & Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA

^d Department of Otolaryngology and Head & Neck Surgery, Vanderbilt University Medical Center, Nashville, TN, USA

^e Vanderbilt Brain Institute, Vanderbilt University Medical Center, Nashville, TN, USA

^f Vanderbilt Kennedy Center, Vanderbilt University Medical Center, Nashville, TN, USA

ARTICLE INFO

Keywords:

'Sensitive period'

Infant speech perception

Syntactic skills

Native Language Magnet theory (NLM)

Developmental speech and language disorders

Individual differences

ABSTRACT

The 'sensitive period' for phonetic learning posits that between 6 and 12 months of age, infants' discrimination of native and nonnative speech sounds diverge. Individual differences in this dynamic processing of speech have been shown to predict later language acquisition up to 30 months of age, using parental surveys. Yet, it is unclear whether infant speech discrimination could predict longer-term language outcome and risk for developmental speech-language disorders, which affect up to 16 % of the population. The current study reports a prospective prediction of speech-language skills at a much later age—6 years-old—from the same children's nonnative speech discrimination at 11 months-old, indexed by MEG mismatch responses. Children's speech-language skills at 6 were comprehensively evaluated by a speech-language pathologist in two ways: individual differences in spoken grammar, and the presence versus absence of speech-language disorders. Results showed that the pre-frontal MEG mismatch response at 11 months not only significantly predicted individual differences in spoken grammar skills at 6 years, but also accurately identified the presence versus absence of speech-language disorders, using a machine-learning classification. These results represent new evidence that advance our theoretical understanding of the neurodevelopmental trajectory of language acquisition and early risk factors for developmental speech-language disorders.

1. Introduction

Decades of research have demonstrated that towards the end of the first year, infants undergo a 'sensitive period' for phonetic learning where their discrimination of native speech contrasts improves while discrimination of nonnative speech contrasts declines (Kuhl et al., 2006; Tsao et al., 2006; Werker and Tees, 1984). The biological and neural mechanisms underlying this important learning process have been extensively examined over the last decade (Kuhl, 2010; Werker and Hensch, 2015).

One central theoretical issue concerns what infants learn and how they learn during this 'sensitive period' and more importantly, how the learning outcome relates to later language development (Best et al., 2016; Kuhl et al., 2005; Kuhl et al., 2008; Werker and Curtin, 2005).

Specifically, the recent Native Language Magnet theory-expanded (NLM-e) (Kuhl et al., 2008) postulates that infant learning of native speech sounds, shaped by input speech characteristics and social interactions, leads to neural commitment to the native language (Native Language Neural Commitment, NLNC) that forges a perception-production link, and is therefore predictive of later language skills. Indeed, Tsao et al. (2004) reported on the first prospective study, showing that 6-month-old native vowel discrimination (using a conditioned head-turn procedure) is *positively* correlated with later language skills at 13, 16, and 24 months of age measured with the MacArthur-Bates Communicative Developmental Inventory (CDI) parental survey (Fenson et al., 1993).

However, a competing explanation is that this predictive relation could be contributed to general auditory processing skills, that is, infants

* Corresponding author at: Institute for Learning and Brain Sciences, University of Washington, Portage Bay Building, Box 357988, Seattle, WA 98195-7988, USA.
E-mail address: zhaotc@uw.edu (T.C. Zhao).

<https://doi.org/10.1016/j.dcn.2021.100949>

Received 29 June 2020; Received in revised form 23 March 2021; Accepted 27 March 2021

Available online 31 March 2021

1878-9293/© 2021 The Authors.

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

with better auditory skills discriminate speech sounds better, and therefore learn languages faster. To address this issue, Kuhl et al. (2005) using behavioral measures, and Kuhl et al. (2008) using brain measures tested infants' *nonnative* speech discrimination in addition to their native speech discrimination skills. If general auditory perceptual skill is the driving cause, then infants' *nonnative* speech discrimination should predict later language skills in a *positive* direction, the same as native speech discrimination. That is, on the general auditory skill explanation, the better an individual infant's *nonnative* speech discrimination is, the better language outcome would be later in development. Alternatively, if neural commitment to native language sounds is the driving cause, then *nonnative* speech discrimination should either be *unpredictive* of later language skills or predict later language in a *negative* direction, the opposite from native speech discrimination. That is, individuals who show better *nonnative* speech discrimination should have worse language outcomes later in development. Kuhl et al. (2005) showed that *nonnative* speech discrimination in fact predicted later language skills, but in the *opposite* direction as native speech discrimination, bolstering the idea of neural commitment. Specifically, infants at 7.5 months of age were tested behaviorally using the head-turn conditioning procedure on both native and *nonnative* consonant discrimination and their language development was followed to 30 months of age using the CDI (Kuhl et al., 2005). The results showed that at 7.5 months of age, 1) native and *nonnative* speech discrimination were negatively correlated, and that 2) individuals with better native speech sound discrimination at 7.5 months showed *faster* language growth, whereas individuals with better *nonnative* speech sound discrimination at 7.5 months showed *slower* language growth. These results were later replicated using EEG-ERP methods to measure native and *nonnative* speech discrimination at 7.5 months at the cortical level (Kuhl et al., 2008). Specifically, the mismatch negativity (MMN) (Naatanen et al., 2007) was used as an index of neural speech discrimination, whereby a larger MMN reflects better neural discrimination. Indeed, the results again demonstrated that individuals with better discrimination of the native contrast, indexed by a larger MMN for the native contrast, showed higher language skills later on while the opposite relation was observed between the *nonnative* contrast and later language skills. These results further demonstrated that better native speech discrimination is reflective of a neural circuitry committed to native language processing, which places the infant on a faster trajectory for native language development, whereas good *nonnative* discrimination is actually a negative predictor – it suggests that infants are not selectively attending to native contrasts, and that infants have uncommitted neural circuitry for native language, thus putting the infants on a slower trajectory for language development.

Following this idea, the current study further expands the theory on two major axes that have not been previously discussed. First, existing studies have only employed a single method of measuring language skills, namely the CDI parental report (Fenson et al., 1993), which is limited to 30 months of age. While the CDI parental report surveys children 8–30 months of age on language skills that are appropriate at different stages (e.g. gestures, word understanding, word production, sentence understanding and sentence production) and has been repeatedly validated, no published studies have examined the long-range predictive effect of infant speech discrimination beyond 30 months of age. Second, while NLM-e focuses largely on typical development and does not make explicit predictions about clinical speech-language disorders, we have a reasonable basis for extending the theory to the atypical realm. Specifically, our prediction is that individuals with over-sensitive (better) *nonnative* speech discrimination predicts a slower language growth trajectory, which could also be an early indicator of speech-language disorders that will call for clinical intervention. Using infant speech perception to predict longitudinal risk for prevalent developmental speech-language disorders, which cumulatively affect up to 16 % of the population but are drastically under-identified (National Academies of Sciences and Medicine, 2016),

could have significant clinical-translational impact.

To address these questions, the current study prospectively followed infants from a previous study (Zhao and Kuhl, 2016) and examined their speech and language skills at a much later age: 6 years of age. As infants, the participants' *nonnative* speech discrimination was measured at 11 months (i.e. at the end of the 'sensitive period'), indexed by the mismatch response measured using Magnetoencephalography (MEG). Here we conducted follow-up tests at 6 years of age: children's speech-language development was comprehensively assessed by a speech-language pathologist (second author OB) in the laboratory setting. The evaluation included standardized assessments of expressive syntactic skills, articulation and speech intelligibility, short-term memory, word reading efficiency, and non-verbal intelligence. We focused on two specific outcomes: 1) individual differences in expressive syntactic skills, measured as a continuous variable on a spoken grammar task and 2) risk of developing speech or language disorders, measured as a binary variable from a comprehensive clinical best estimate of presence versus absence of disordered speech or language.

Higher sensitivity to *nonnative* speech stimuli has previously been shown to predict lower sentence complexity and mean length of utterance (MLU) at 24 and 30 months respectively; this is important because sentence complexity and MLU are both early markers of grammatical competency in children (Kuhl et al., 2008). The current study sought to examine whether the previously observed relationship between speech discrimination and expressive grammar outcomes at 24 and 30 months extends later into childhood and formal schooling (i.e., at 6 years of age). At the same time, expressive syntactic skill (i.e., spoken grammar ability) in early elementary school is a crucial component of overall language competence. It contributes to school readiness, reading comprehension, and social skills among peers (Brimo et al., 2017; Catts et al., 2002; Fujiki et al., 1999). Skills such as following complex directions in a classroom, storytelling, peer collaboration, and early reading all rely on a foundation of efficient processing of syntactic structures. Moreover, communication barriers related to poor syntactic competence contribute to negative downstream effects in children's academic success and social-emotional development (Hubert-Dibon et al., 2016). Therefore, expressive syntactic skill was chosen as our primary measure of interest.

In order to test whether the NLM-e theory can generalize and generate predictions in a clinical context, our second outcome of interest uses comprehensive clinical assessment to identify children with atypical speech-language development (i.e., presence of speech-language disorder). Our criteria for atypical speech-language development at 6 years of age included clinically classifiable articulation difficulties or language structure problems. These broad criteria follow the *trans-diagnostic* approaches that target shared underlying mechanisms across disorders that may present as clinically distinct (Sauer-Zavala, 2017). If infant *nonnative* speech discrimination turns out to be a reliable predictor of atypical speech-language development, it would open up avenues for robust prediction of individual risk of developmental speech-language disorders at a much earlier age than currently thought possible. Understanding these risk factors will ultimately allow the improvement of early identification and access to early intervention, following an epidemiological framework (National Academies of Sciences and Medicine, 2016; Raghavan et al., 2018).

Therefore, the current study tested two specific questions: 1) Can *nonnative* speech discrimination at 11 months of age, indexed by MEG-measured mismatch response, predict expressive syntactic skills at age 6? And 2) Can infants' *nonnative* mismatch response also help accurately detect the presence of speech-language disorders at age 6, thus providing a window on atypical versus typical speech-language development? In addition, the original mismatch response measured by MEG also allows examination of the hypothesized effects from two specific cortical regions separately, namely prefrontal and temporal regions of the brain that are largely thought of as the neural generators of the mismatch response (Alho, 1995).

2. Method

2.1. Participants

Participants were 27 typically developing children being raised in monolingual English-speaking households (mean age = 6.12 years, min = 5.91 years, max = 6.35 years, std = 0.14 years). All children previously participated in the Zhao and Kuhl (2016) study as infants. Out of the original 38 infants, 2 were excluded from the current study due to a diagnosis of significant developmental disorders not compatible with our test battery (e.g. autism), 9 were either not reached or could not participate (e.g. moved out of state, did not arrive at the appointment), 27 returned for the in-laboratory behavioral testing, with 1 child who could not complete the testing. Out of the 26 children who completed standardized language tests at 6 years of age, 23 also had usable MEG mismatch response for nonnative speech discrimination at 11 months of age that was included in the initial publication. Note that 10 of the 23 individuals were in the original music intervention group.

2.2. Neural speech discrimination at 11 months

The original infant MEG experiment (Zhao and Kuhl, 2016) utilized a nonnative speech contrast based on a consonant duration change embedded in disyllabic nonwords (/bibbi/ vs. /bibi/). On 85 % of the trials, infants were presented the disyllabic nonword with a long consonant between the vowels (i.e., /bibbi/); in the 15 % deviant trials, the syllable structure was violated by shortening the length of the middle consonant from 150 ms to 50 ms (i.e., /bibi/) (Fig. 1A), effectively reducing the silent gap between two vowels by 100 ms. This difference reflects an acoustic feature for phonemic contrasts in languages such as Japanese and Finnish, but not used in English (Aoyama, 2000). We also

separately recorded the neural response to /bibi/ when it was presented in a constant stream (as standard for 200 trials); we subtracted neural responses to /bibi/ when it served as standard from neural responses to /bibi/ when it served as deviant (200 trials) in the context of the syllable /bibbi/. This design removes neural responses to acoustic differences between the standard and the deviant (Kujala et al., 2001). The MMR time window (Fig. 1B, shaded region) was timed to the onset violation (i.e., onset of the 2nd syllable, Fig. 1A).

2.3. Speech-language evaluation at 6 years old

Children completed a comprehensive battery of speech and language assessments (described below) administered by a speech-language pathologist (second author OB) during a single visit to the laboratory (approximately 1.5 h), with breaks as needed. The Speech-Language Pathologist and other research staff involved in testing were blind to participants' individual MEG data at the time of 6-year-old visit and during the presence-versus absence classification process (described below). Parents completed a demographic questionnaire describing parent formal education, child's medical and developmental history, parent speech and language concerns, and parent musical experience. All procedures were approved by IRBs at both the University of Washington and Vanderbilt University Medical Center. Informed consents were obtained from all parents and assents were obtained from all participating children prior to the testing procedure. All participating families were compensated monetarily.

The comprehensive battery of speech and language assessment took place in a quiet booth and included tests as follows: 1) a bilateral pure-tone hearing screening (20 dB SPL at 1000, 2000, and 4000 Hz) was first administered to confirm hearing within functional limits. 2) The Test of Early Grammatical Impairment Screening Probe (TEGI; (Rice and

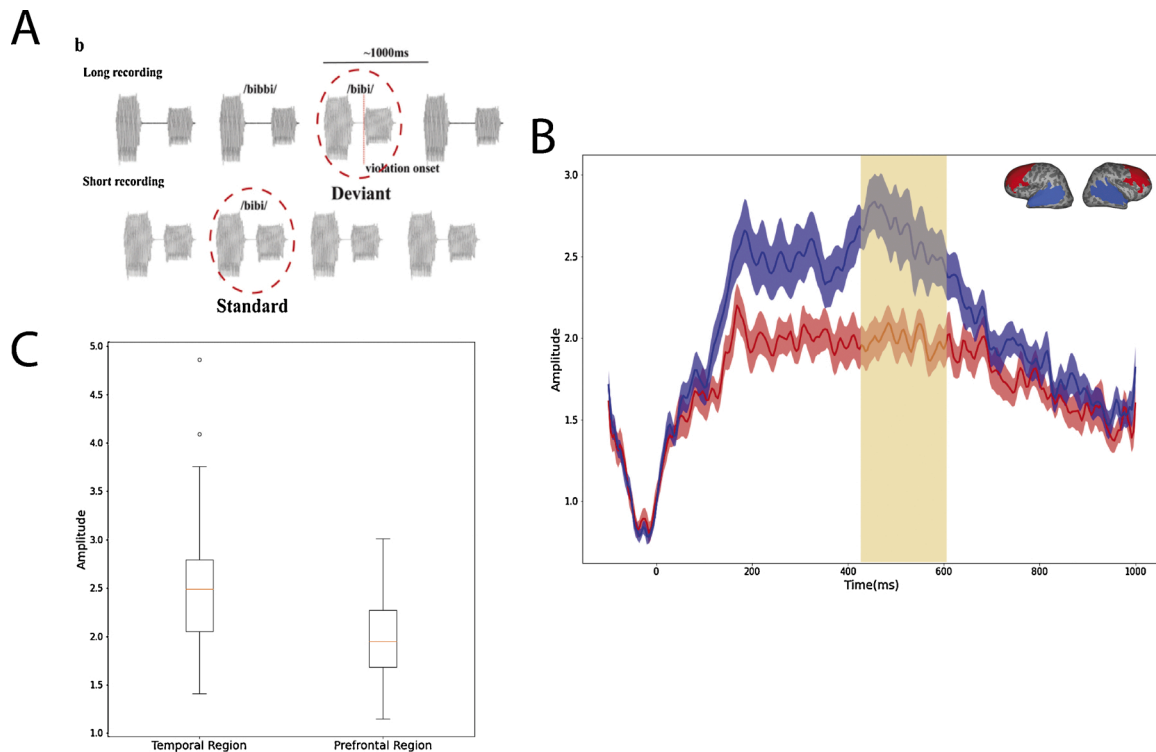


Fig. 1. A) Schematics of MEG experiment: in the long recording, a traditional oddball paradigm was used where deviants /bibi/ was presented 15 % of the time, among the standards /bibbi/ (85 % of the time). In a separate shorting recording, the same number of /bibi/ was presented in a constant stream as standard. MMR was calculated as the difference between identical stimulus (i.e. /bibi/) when presented as standard vs. as deviant. B) Group average MMR for the 23 participants in the current study in the temporal (blue) and prefrontal (red) regions. The time window for MMR is shaded in yellow. C) Box plot of MMR (averaged across the shaded time window) for Temporal and Prefrontal region. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Wexler, 2001)) was administered both to rule out phonological disorders that would impact validity of language testing and to screen for language impairment. Children were asked to answer questions about various pictures designed to elicit specific phonological sounds and morphological endings such as past tense and third person singular, and were assessed to be either at/above or below criterion for the TEGI Screening Probe. 3) The Primary Test of Nonverbal Intelligence (PTONI; Ehrler and McGhee, 2008)) was administered to assess and control for age-normed nonverbal IQ. Children were shown a series of images and asked to point to the object that did not belong. Results from the PTONI were quantified through standard scores. 4) The Number Recall subtest of the Kaufman Assessment Battery for Children, Second Edition (KAB-C-II; Kaufman and Kaufman, 2004)) was used to assess short-term memory. Raw scores from this subtest were used to quantify short-term memory performance. 5) The Structured Photographic Expressive Language Test, 3rd Edition (SPELT-3; Dawson et al., 2003)) was used as our primary measure of expressive syntax, and was quantified as a standard score. Children were asked to describe pictures that were designed to elicit particular morphosyntactic forms, and their responses were scored based on appropriate use of those forms. 6) The Sentence Imitation subtest of the Test of Language Development – Primary: 4th Edition (TOLD-P:4; Newcomer and Hammill, 2008)) was also used to assess expressive language. A scaled score was derived from this subtest. Sentence imitation tasks have been shown to be reliable and sensitive in identifying children with language impairment (Archibald and Joanisse, 2009; Redmond et al., 2011). 7) Speech sound errors observed throughout testing and conversation with the participant were used to flag errors that were inappropriate for age and gender. The Sounds-in-Words subtest of the Goldman-Fristoe Test of Articulation, 3rd Edition (GFTA-3; Goldman and Fristoe, 2015) was administered and a standard score was derived to assess speech and articulation in participants flagged by the clinician for speech-sound concerns. And finally, 8) the Test of Word Reading Efficiency – Second Edition (TOWRE-2; Torgesen et al., 2012)) was administered to assess fluency and accuracy of both sight word and nonword reading. Children were given lists of sight words and nonwords and asked to read words aloud as fast and accurately as possible. Based on rate and accuracy of children's responses, standard scores for sight word and non-word reading were derived. The TOWRE-2 is normed for ages 6 years 0 months to 24 years and is not appropriate for children who are not yet reading; therefore 16 out of 26 children could not complete this assessment ($n = 4$ had not yet reached their sixth birthday on the date of testing, and of the 22 children who were old enough to complete the TOWRE, testing was discontinued for 12 of them based on inability to complete the practice items).

Each participant was assigned to one of three outcome types based a combination of history of speech-language impairment, standardized assessment scores, and clinical observation during the testing session: A) typical speech-language development: absence of speech and/or language disorder; B) mild concern for speech-language disorder; or C) atypical speech-language development: presence of speech and/or language disorder. Participants were placed into the 'presence of speech and/or language disorder' category if they had a reported history of speech-language disorder or intervention, or one or more standardized speech or language assessment scores that were more than 1.5 standard deviations below the mean. Based on these criteria, 5 out of the 26 tested children were assigned to this category. Two of these 5 children had a history of speech-language disorder which had since resolved: one child received services related to articulation difficulties, while one child experienced stuttering from ages 2 through 5. While both of these children were classified as having atypical speech-language development, neither were included in the final analyses as their MEG data was not usable/available. One additional child had an active individualized education plan (IEP) for speech-language services related to receptive language skills. The remaining 2 children showed atypical articulation, with GFTA-3 standard scores more than 1.5 standard deviations below the mean. All clinician concerns were shared with parents at the time of

the visit, who were encouraged to follow up with their child's pediatrician and teachers.

Participants were placed in the 'mild concern' category if there was concern for speech and language difficulties that did not meet the threshold for therapeutic intervention. The 'mild concern' category (B) required a combination of 1) clinician observations related to language use, articulation, pre-literacy performance from the TOWRE-2, or other atypical features such as disfluency or pragmatic oddities, and 2) parent concerns related to speech, language, or academic performance. Only 1 child was assigned to this category based on parent concern in multiple speech and language areas, and clinical observation related to atypical pre-literacy skills. All other participants were determined to have an absence of any speech or language disorder and therefore classified in the typical speech-language category.

In all further analyses, we merged the 'mild concern' category with the 'atypical speech-language' category.

For the final sample of 23 children who had both neural and behavioral data available, two key outcomes were derived from the behavioral session: 1) expressive syntactic skills measured by the Structured Photographic Expressive Language Test, 3rd Edition (SPELT-3); and 2) classification of presence of speech-language disorders (atypical) vs. absence of speech-language disorders (typical speech-language development) based on the comprehensive clinical assessment (see Table 1 and Table 2 for participant classification summary).

2.4. Data analyse

The final dataset (OSF link) consisted of 23 individuals with both neural speech discrimination at 11 months of age and behavioral speech-language assessment at 6 years of age.

Regression analyses were conducted to address our first research question: can infants' mismatch response to the nonnative speech contrast at 11 months predict individual differences in spoken grammar skills (age-normed SPELT-3 standard scores) at 6 years of age. Specifically, we conducted parametric regression analyses as well as a machine learning support-vector regression. *Parametric*: We examined the relation between mismatch response and SPELT-3 while controlling for nonverbal intelligence (PTONI). The mismatch response values were the published values averaged across the corresponding time windows for each region (Fig. 1B, C) (Zhao and Kuhl, 2016). We conducted separate regression analyses for the prefrontal region mismatch response and the temporal region mismatch response. In addition, to rule out potential effects from the original intervention, we conducted an additional regression analysis between mismatch response and SPELT-3 while the original group assignment (music intervention vs. control) was also controlled for in the model. All parametric regression analyses were done using SPSS Version 19 (IBM). *Support-vector regression*: We utilized a machine-learning support-vector regression (SVR) to further examine

Table 1

Results of speech-language evaluation at age 6, data presented as mean (\pm standard deviation) unless otherwise stated, for the 23 children included in the final dataset.

	Typical language category (absence of speech-language disorder)	Atypical speech-language category (presence of speech-language disorder)
Number of Participants	19 (13 males)	4 (3 males)
Age	6.10 (\pm 0.15)	6.15 (\pm 0.10)
Median SES	College/Technical Degree (3–4 years)	College/Technical Degree (3–4 years)
Non-verbal IQ (PTONI)	113.26 (\pm 21.25)	113.8 (\pm 21.0)
Expressive grammar (SPELT-3)	113.37 (\pm 4.23)	112.8 (\pm 5.56)

Table 2

Description of classification criteria met for inclusion in atypical speech-language category.

Participant	Classification criteria met:
S1	Currently receiving speech-language services through school system related to receptive language and pre-reading skills, per parent report.
S13	Speech sound delay as demonstrated by standard score of 73 (>1.5 SD) on Goldman Fristoe Test of Articulation.
S16	Child could not complete Test of Word Reading Efficiency practice items due to difficulty recognizing letters. Parental concern for speech-language disorder based on survey with pre-reading skills (child has difficulty recognizing letters and sound to letter correspondence). Note that parent also reported co-occurring sensory integration challenges.
S26	Speech sound delay as demonstrated by standard score of 62 (>2 SD) on Goldman Fristoe Test of Articulation.

the relation between infant mismatch response and expressive synthetic skills at age 6 (Drucker et al., 1996). The model uses whole mismatch response time series, therefore takes into consideration the temporal dynamics of the mismatch response at prefrontal and temporal regions. All machine-learning analyses for this current study were done using the open source scikit-learn package (Pedregosa et al., 2011). The dataset is first randomly split into a training and a testing set. Mismatch response time series (instead of averaged values) in the training set were first used to fit the model with a linear kernel function ($C = 1.0$, $\epsilon = 0.1$). Once the model is trained, the mismatch response time series from the testing set were then used to generate predictions of the SPELT scores. A k-fold cross-validation method ($K = 2$) was also used to enhance model prediction. The R^2 coefficient of determination between actual measured SPELT and model predicted SPELT is taken as index of model performance.

To further evaluate the model performance, we shuffled the SPELT scores across individuals and then conducted the same SVR analyses. In such cases, the mismatch response time series should bear no predictive value to SPELT score and the R^2 should reflect a model performing at chance level. We repeated this process 1000 times and generated an empirical null distribution of R^2 and we compared our originally obtained R^2 coefficient against this distribution (Xie et al., 2019).

To address our second research question of whether we could detect risk for developing speech and language disorders using infant mismatch response time series, we utilized a supervised Support-Vector Machine (SVM) classification method (Cortes and Vapnik, 1995). Similar to the SVR analyses, the SVM takes all the time points of the mismatch response time series into the model. However, instead of predicting a continuous variable (e.g. SPELT score), the SVM generates a binary categorical prediction (i.e. typical vs. atypical speech-language categories). A nested cross-validation was adopted to enhance model performance while addressing the issue of bias related to having a small sample size for SVM (Vabalas et al., 2019). The dataset was first split into two folds that preserves the percentage of samples for each class (i.e. typical vs. atypical). Then each fold served first as training set and then as testing set, or vice versa. Two separate SVM models with linear kernel ($C = 1.0$, $\epsilon = 0.1$) were developed based on the two training sets and were then tested on the two testing sets. That is, based on the trained model, testing set mismatch response time series was used to generate a prediction of presence versus absence of speech and language disorders. The prediction accuracy score (percentage of correct labels) in the testing for each model was first calculated and the overall accuracy was calculated as the mean across the two models. The overall true positive rate (TPR) /sensitivity and overall true negative rate (TNR) /specificity were also calculated (TPR = True Positive/(True Positive + False Negative); TNR = True Negative / (True Negative + False Positive)). Similarly, an empirical null distribution of accuracy scores was built by shuffling the category labels across individuals and repeating SVM with nested cross-validation 1000 times. We then compared the obtained accuracy score to the empirical null distribution.

3. Results

3.1. Infant speech discrimination as a predictor of individual differences in spoken grammar

Results from the parametric regression analyses suggested that the prefrontal region mismatch response at 11 months significantly predicted SPELT-3 scores at 6 years of age while controlling for non-verbal intelligence ($R^2 = 0.35$, $p = 0.026$, Beta (mismatch response) = -0.56 , $p = 0.014$, Bonferroni correction applied) (Fig. 2A, left column). The results remain in the additional model where the original group assignment was controlled ($R^2 = 0.369$, $p = 0.03$, Beta (mismatch response) = -0.546 , $p = 0.009$). In contrast to the prefrontal results, results showed that the temporal region mismatch response at 11 months did not predict SPELT-3 scores at 6 years of age ($R^2 = 0.19$, $p = 0.254$, Bonferroni correction applied) (Fig. 2B, left column).

The SVR analyses suggested that when using the whole prefrontal region mismatch response time series, the R^2 between predicted and measured SPELT-3 is 0.305, higher than the 97.5th percentile of empirical null distribution for R^2 (i.e., $R^2 = 0.157$) (Fig. 2A, middle and right columns), in line with results of parametric analyses. However, when using the whole temporal region mismatch response time series for SVR analyses, the R^2 between predicted and measured SPELT-3 is 0.108, still slightly better than the 97.5th percentile of the null distribution (i.e., $R^2 = 0.009$) (Fig. 2B, middle and right columns). In this case, using the time series of the temporal region mismatch responses achieved better-than-chance prediction of SPELT-3 score.

To summarize, both parametric and machine-learning regressions demonstrate that we can predict individual expressive syntactic skills at 6 years of age using mismatch responses to nonnative speech contrasts at 11 months of age. SVR overall performs better than parametric regression analyses, suggesting potentially more sensitive predictions when taking into consideration the temporal dynamics of the time series. More importantly, this predictive relation is stronger when the prefrontal region mismatch response is used.

3.2. Infant speech discrimination as a predictor of presence of speech-language disorders

Results from the SVM classification analyses revealed a similar pattern of results. Using the whole prefrontal region mismatch response time series, an accuracy score of 86.74 % was achieved to classify participants into 'typical' vs. 'atypical' categories. This is at the 97.5th percentile of the empirical null distribution of accuracy scores created by shuffling the labels ('presence of speech or language disorder - atypical' vs. 'absence of speech or language disorder - typical') across participants (Fig. 3A). In addition, the TPR and TNR are 0.5 and 0.94, respectively. However, when using the temporal region mismatch response time series, the accuracy score was only 82.57 %, below the 97.5th percentile of the random distribution (Fig. 3B). The TPR and TNR are 0 and 1.0 with no true positive identified. Together, the results demonstrate that prefrontal MMR, but not temporal MMR, can predict, at a higher than chance level, whether an infant will be likely to manifest atypical speech-language development prior to age 6.

4. Discussion

The Native Language Magnet theory expanded (NLM-e) predicts that infant speech discrimination during the 'sensitive period' for phonetic learning predict later language skills, through the mechanism of native language neural commitment (NLNC) to the native language. According to this model, discrimination of nonnative phonetic contrast is an especially sensitive predictor of slower language development (Kuhl et al., 2005, 2008). We followed this theoretical framework and extended the current literature on two important axes: 1) we demonstrated that the mismatch response indexing nonnative speech

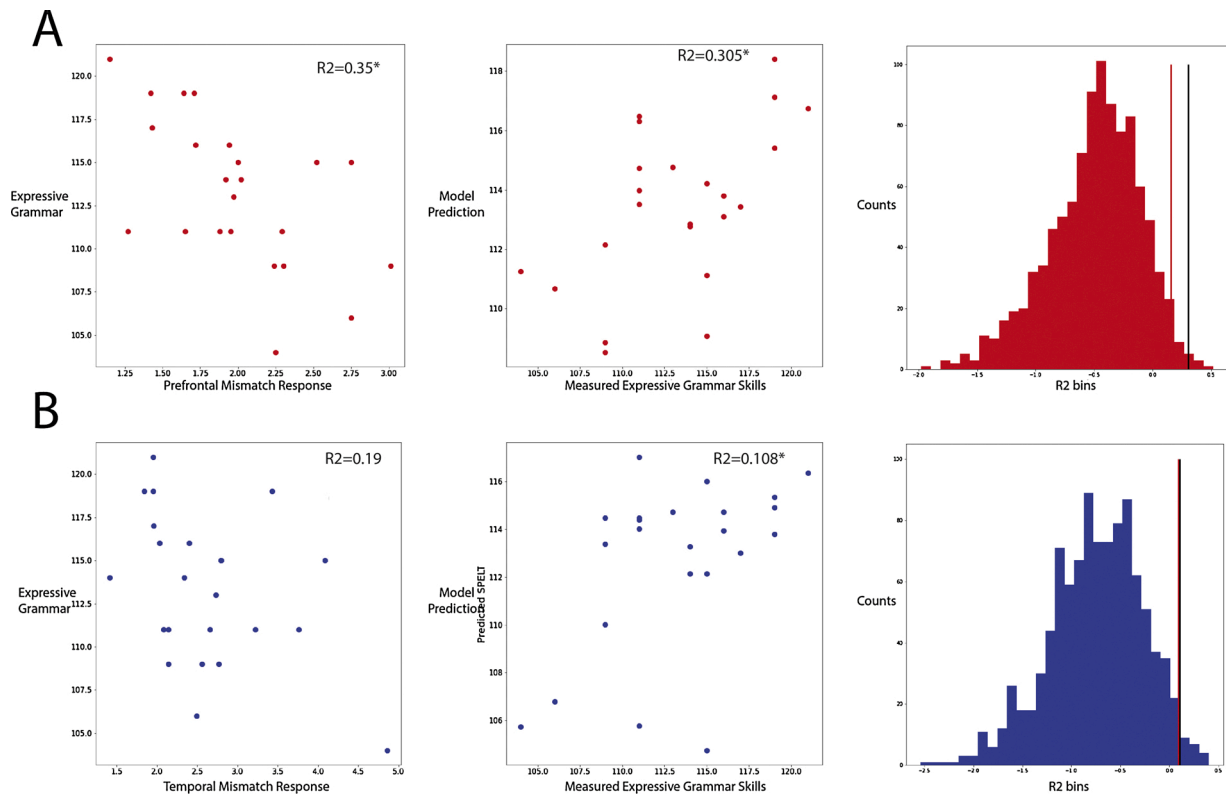


Fig. 2. A) Left: scatter plot of averaged mismatch response in prefrontal region and expressive grammar skills (measured by SPELT-3). Middle: Measured expressive grammar scores and SVR model predicted expressive grammar scores from the whole prefrontal mismatch response time series. Right: Empirical null distribution of R2. Black line: R2 from the current dataset. Red line: 97.5th percentile in the distribution. B) Left: scatter plot of averaged mismatch response in temporal region and expressive grammar skills. Middle: Measured expressive grammar scores and SVR model predicted expressive grammar scores from the whole prefrontal mismatch response time series. Right: Empirical null distribution of R2. Black line: R2 from the current dataset. Red line: 97.5th percentile in the distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

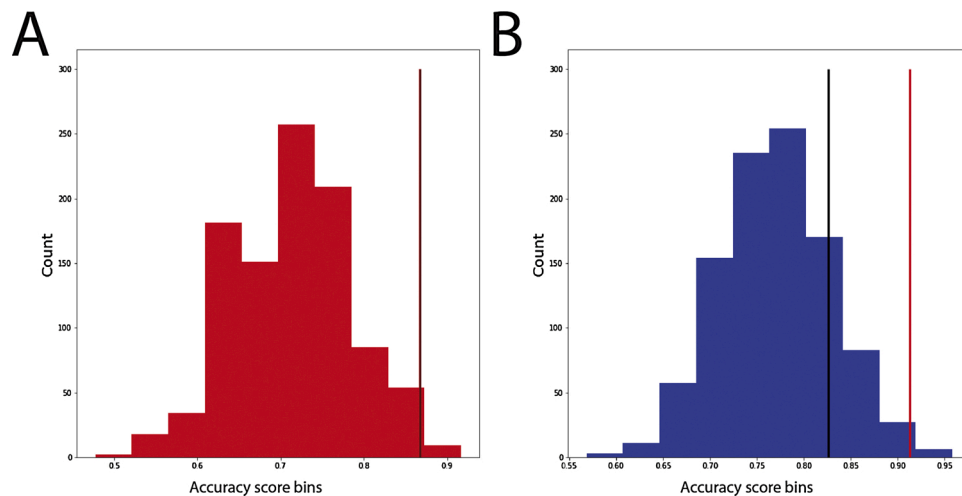


Fig. 3. A) Random distribution of accuracy scores from SVM classification, using prefrontal MMR. The current data performance is at 97.5th percentile. B) Random distribution of accuracy scores from SVM classification, using Temporal MMR. The current data performance (black line) is significantly below 97.5th percentile.

discrimination at 11 months of age prospectively predicted individual differences in spoken grammar skills at a much later point in development, at 6 years of age. 2) We demonstrated that the infant nonnative speech discrimination can generalize and generate predictions of clinical relevance; that is, using a machine-learning approach, the model accurately identified which infants would later exhibit presence of speech-language disorders, providing a window on atypical speech-language development. Moreover, the MEG data showed that the

brain area best predicting language development is largely localized to the prefrontal region; the auditory region did not predict long-term language development. Taken together, the current results are consistent with prior findings obtained from parent report of expressive language skills at 2.5 years of age (Kuhl et al., 2008; Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; Tsao et al., 2004); and further, the study expands the theoretical framework of NLM-e significantly, demonstrating that infant speech discrimination is one of the earliest milestones that sets

individuals on their speech-language acquisition trajectory.

4.1. Early predictors of later language skills

Other early predictors of later language skills have also been reported. In the speech domain, using retrospective methods, Newman and colleagues (Newman et al., 2006) reported that infants' ability to segment speech before 12 months of age is related to later language skills measured by both parent survey at 24 months of age and direct language assessment between 3–4 years of age. A series of other work shows that word processing speed at age 25 months is a significant predictor of expressive language skills at 8 years of age (Fernald et al., 2006; Marchman and Fernald, 2008). More recently, researchers have also reported that environmental factors, such as maternal speech input quality, including percent of infant- or child-directed speech, the amount of one-on-one interaction, and lexical properties in the speech, are important contributing factors to later language skills (Ferjan Ramírez et al., 2019, 2020); Newman et al., 2016; Ramírez-Esparza et al., 2014).

In addition to speech predictors, other work has demonstrated that general auditory processing skills can also predict later language skills as well as detect language disorders (Benasich and Tallal, 2002; Molfese, 2000; Molfese and Molfese, 1997). A seminal study showed that neonates' obligatory evoked responses to speech sounds predicted language skills in the same children when they turn 3 years of age (Molfese and Molfese, 1985). Using another approach, researchers have shown that the ability to process rapid auditory cues is predictive of later language skills, and it distinguishes children at-risk of language disorders from those with typical language (Benasich and Tallal, 2002). More recently, researchers described neural signatures related to rapid auditory processing in infants that are related to later language skills and also differentiate children at-risk for language impairment (Cantiani et al., 2019). Recent research also suggests that the ability to process music rhythm can be related to language skills, and these authors hypothesized that musical rhythm is a potential early predictor of language skills and developmental speech and language disorders (Gordon et al., 2015).

Thus far, the mechanisms by which these predictors (native speech vs. nonnative speech vs. nonspeech/general auditory) are interconnected and complement each other remains unknown. Future research in larger samples is warranted to systematically examine the interconnections, interactions and relative contributions of different predictors with the aim of identifying an ideal set of measures in infants that can generate best predictions for future language skills and also produce best detection of at-risk infants for early intervention. To illustrate, in our current study we utilized only a single nonnative speech contrast to generate prediction, we did not incorporate any native speech contrast or nonspeech contrast. Future studies will need to test a variety of native and nonnative speech contrasts that utilize a variety of cues to replicate this result and characterize potentially different predictivity from different speech contrasts. Such study will also allow further examination of different theoretical perspectives with regard to the underlying mechanisms through which native vs. nonnative speech contrasts drive such predictions. In addition, it would also be important to systematically compare speech processing with a range of tests assessing infants' general auditory skills, especially ones related to temporal information processing, to examine their long-term predictive value for children's language skills.

4.2. Clinical application for mitigating risk of speech-language disorders

The current study provides an important step for future work into the predictive power of speech processing in infancy and speech-language outcomes in childhood. Given the high prevalence of developmental speech and language disorders and their academic, social-emotional, financial, and vocational impacts later in life, early identification and preventive intervention in infants at risk for communication

impairments could mitigate these consequences in childhood and beyond (NASEM, 2016). Preventive approaches for very early identification and treatment of children at risk of communication disorders have already shown promise in improving child outcomes, such as in a pilot study for infants with galactosemia, which is frequently accompanied by expressive language impairments (Peter et al., 2020). The potential for noninvasive approaches such as MEG to assist in identifying children with speech discrimination characteristics that put them at greater risk of speech-language disorders, is worthy of expansion and further exploration as means to make a significantly positive impact on public health.

4.3. Language assessment in children

While all children in this study were comprehensively evaluated for identifying atypical speech-language development, expressive grammar itself could also be a particularly critical measure that can indicate speech and language disorders. Children with language impairment have notable weaknesses in morphology and syntax (Owen and Leonard, 2006). In addition, children with late language emergence (often referred to as "late-talkers") also have greater vulnerability for continued weaknesses in grammar later in childhood, compared to other domains of language such as vocabulary (Rice et al., 2008). In our current sample, children with typical vs. atypical speech-language development did not differ on their expressive grammar scores. Future studies in larger samples should investigate whether poor grammar skill is a main driving mechanism for language deficits in children using a large sample and can investigate associations/dissociations between phonological and grammatical skill in relation to presence or absence of developmental speech/language disorders.

On the other hand, our current results provide indirect support for continuity of measures for syntactic skills over development. The Sentence Complexity and the Mean of Longest Intelligible Utterances (M3L) metrics from the CDI have strong concurrent validity with metrics of grammatical complexity from behavioral language sampling in both children with typical development and language impairment (Law and Roy, 2008; Thal et al., 1999; Dale, 1991). However, their predictive validity past the third year of life warrants further investigation. Indeed, our current work, along with prior work indirectly support such predictive validity given that infant nonnative speech discrimination similarly predicts CDI measures at 30 months and formal assessment of expressive grammar at 6 years.

4.4. Machine-learning methods as a powerful tool

In the future quest to understand the interconnections of large sets of infant characters and their combined predictive values of later language skills, automated machine-learning methods are becoming especially relevant and useful. Indeed, there is an increasing amount of interest in applying machine learning methods to identify individuals affected or at-risk for developmental communication disorders, such as autism (Bosl et al., 2011; Justice et al., 2019; Zare et al., 2016).

Our current results complement the existing efforts and suggest that we can leverage machine-learning methods to identify healthy and typically developing infants who will later develop speech-language disorders in the mild-moderate range. Given the growing evidence that language impairments occur on a continuum rather than dichotomously (Lancaster and Camarata, 2019), similarly to mental health traits in the population (Martin et al., 2018), it becomes even more important to both identify biological mechanisms underlying individual differences in language acquisition, and uncover prodromal markers of later clinical speech or language problems that might be relatively mild but still require intervention to mitigate long-term academic and social consequences. While machine learning methods can be powerful, the caveats of applying ML models on neural data is that there is generally a very small sample in each individual dataset such that the inherent bias

in these samples may be amplified by the ML models (Vabalas et al., 2019). Future research is warranted in which large longitudinal population-based datasets can be accessed or aggregated to allow us to further examine differences across a wider continuum of speech and language impairment. More complex machine-learning models can then be applied to more accurately predict children's language outcome from the earliest ages with less bias, with the ultimate goal to intervene during early development when they are the most effective.

4.5. The role of inferior frontal gyrus

The importance of the prefrontal mismatch response in its predictive power for later language skills over the temporal mismatch response is of particular interest. It leads to some interpretation and speculation regarding how such long-range predictions are possible, considering the important role of the inferior frontal gyrus (IFG) in both phonetic learning and syntax processing. The IFG, which overlaps with Broca's area, is demonstrably crucial to speech perception and production (Hickok and Poeppel, 2007). More recently, the advancement in infant MEG neuroimaging has allowed us to directly observe the change in IFG during phonetic learning (Imada et al., 2006; Kuhl et al., 2014). Specifically, pre-verbal young infants are engaging motor areas (e.g., IFG, Cerebellum, see Kuhl et al., 2014) of the brain as they listen to speech, long before they can actually produce speech; and such engagement is interpreted as infants' attempting to simulate the motor gestures required to produce the speech sounds they hear. In these studies, infants' brain activity in IFG increased for nonnative contrasts between 6 and 12 months of age, suggesting that nonnative contrasts become too difficult to simulate in comparison to native speech sounds and thus recruit additional neural resources. In the current study, infants who had higher prefrontal region mismatch (including IFG) are likely to be more reactive to nonnative speech contrasts, thus showing a pattern of less committed brain structure to native language contrasts at 11 months, which appears to have placed them on a slower language development trajectory, evidenced by their outcomes at age 6. These findings suggest that the functional and structural development of the inferior frontal gyrus may have an important role in the 'sensitive period' for phonetic learning (see Kuhl, 2021 for additional discussion).

The link between infant speech discrimination in the prefrontal regions and school-age syntactic skills may also implicate the IFG's central role in hierarchical processing, which manifests as syntactic (grammatical) task performance in language (Fitch and Martins, 2014; Friederici, 2020; Heard and Lee, 2020). Activation of the IFG is present during syntactic encoding in adults even in the absence of lexical/semantic information, and the increase in activation magnitude correlates with increases in syntactic complexity (Pallier et al., 2011). In particular, the pars triangularis is activated in a distinct temporal sequence across lexical, inflectional, and phonological expressive language tasks (Sahin et al., 2009). In addition, the left IFG's role is not limited to processing syntactic information; rather it works in concert with the left posterior middle temporal gyrus and the bilateral supplementary motor area to support many aspects of receptive and expressive language (Segaert et al., 2011). As children develop, the resting state functional connectivity of these regions becomes increasingly lateralized (Xiao et al., 2016) and activation of Broca's area during sentence processing shifts from the opercular IFG to the pars triangularis (Visiennon et al., 2017). In children and adults with speech sound (articulation) disorders, the left IFG shows hyperactivation during task performance compared to TD controls, suggesting that individuals with speech sound disorders may rely on the left IFG as a compensatory network for impairments in the phonological processing loop (Tkach et al., 2011; Preston et al., 2012; Liégeois et al., 2014). As knowledge evolves regarding the complex structural and functional divisions within Broca's area (Fedorenko and Blank, 2020), future research should systematically examine the developmental trajectory of the IFG in relation to infant phonetic processing and later syntactic perception and

production in older children and adults, along with risk of developing disordered speech or language.

5. Conclusion

We reported on a theory-driven prospective follow-up study and demonstrated a robust prediction of 6-year old's language skills from the same child's nonnative speech discrimination at 11 months of age, indexed by neural mismatch response measured with MEG. Specifically, using both parametric statistics and machine-learning approaches, we demonstrated that the prefrontal region mismatch response not only predicted the individual differences in expressive syntactic skills but also detected risk of developmental speech-language disorders with high accuracy. These results are promising in their contribution to the current literature of potential robust early predictors of clinically-relevant language skills in childhood.

Statement for conflict of interests

All authors declare no conflict of interests.

Data Statement

As indicated in the manuscript, the data reported here consisted of 23 individuals with both neural speech discrimination at 11 months of age and behavioral speech-language assessment at 6 years of age and it can be accessed here (OSF link).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This project was supported by funding from the National Institutes of Health, NIH Common Fund through the Office of Strategic Coordination/Office of the NIH Director under Award Number DP2HD098859, and the I-LABS Research on Infants, Children, and Adolescents Grant. Use of REDCap was made possible by Grant UL1 TR000445 from the National Center for Advancing Translational Sciences (NCATS)/NIH. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank our research assistants at I-LABS to make this study possible and all the participating families for their devotion to science.

References

- Alho, K., 1995. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNM) elicited by sound changes. *Ear Hear.* 16 (1), 38–51. <https://doi.org/10.1097/00003446-199502000-00004>.
- Aoyama, K., 2000. A psycholinguistic perspective on Finnish and Japanese prosody: perception, production and child acquisition of consonantal quantity distinctions. *DAIA Dissertation Abstracts International, Section A: The Humanities and Social Sciences* 61 (4), 1376.
- Archibald, L., Joannisse, M., 2009. On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *J. Speech Lang. Hear. Res.* 52 (4), 899–914.
- Benasich, A.A., Tallal, P., 2002. Infant discrimination of rapid auditory cues predicts later language impairment. *Behav. Brain Res.* 136 (1), 31–49. [https://doi.org/10.1016/s0166-4328\(02\)00098-0](https://doi.org/10.1016/s0166-4328(02)00098-0).
- Best, C.T., Goldstein, L.M., Nam, H., Tyler, M.D., 2016. Articulating what infants attune to in native speech. *Ecol. Psychol.* 28 (4), 216–261. <https://doi.org/10.1080/10407413.2016.1230372>.
- Bosl, W., Tierney, A., Tager-Flusberg, H., Nelson, C., 2011. EEG complexity as a biomarker for autism spectrum disorder risk. *BMC Med.* 9 (1), 18. <https://doi.org/10.1186/1741-7015-9-18>.
- Brimo, D., Apel, K., Fountain, T., 2017. Examining the contributions of syntactic awareness and syntactic knowledge to reading comprehension. *J. Res. Read.* 40 (1), 57–74. <https://doi.org/10.1111/1467-9817.12050>.
- Cantiani, C., Ortiz-Mantilla, S., Riva, V., Piazza, C., Bettoni, R., Musacchia, G., et al., 2019. Reduced left-lateralized pattern of event-related EEG oscillations in infants at

- familial risk for language and learning impairment. *NeuroImage: Clinical* 22, 101778. <https://doi.org/10.1016/j.nicl.2019.101778>.
- Catts, H.W., Fey, M.E., Tomblin, J.B., Zhang, X., 2002. A longitudinal investigation of reading outcomes in children with language impairments. *J. Speech Lang. Hear. Res.* 45 (6), 1142–1157. [https://doi.org/10.1044/1092-4388\(2002\)093](https://doi.org/10.1044/1092-4388(2002)093).
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20 (3), 273–297. <https://doi.org/10.1007/bf00994018>.
- Dale, P.S., 1991. The validity of a parent report measure of vocabulary and syntax at 24 months. *J. Speech Hear. Res.* 34 (3), 565–571. <https://doi.org/10.1044/jshr.3403.565>.
- Dawson, J.I., Stout, C.E., Eyer, J.A., 2003. *SPELT-3: Structured Photographic Expressive Language Test*. Janelle Publications, Incorporated.
- Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V., 1996. Support vector regression machines. In: *Paper Presented at the Proceedings of the 9th International Conference on Neural Information Processing Systems*. Denver, Colorado.
- Ehrler, D.J., McGhee, R.L., 2008. *Primary Test of Nonverbal Intelligence*. PRO-ED, Austin, TX.
- Fedorenko, E., Blank, I.A., 2020. Broca's Area Is not a natural kind. *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2020.01.001>.
- Fenson, L., Dale, P., Reznick, J.S., Thal, D., Bates, E., Hartung, J., et al., 1993. *MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. Singular Publishing Group, San Diego, CA.
- Ferjan Ramirez, N., Lytle, S.R., Fish, M., Kuhl, P.K., 2019. Parent coaching at 6 and 10 months improves language outcomes at 14 months: a randomized controlled trial. *Dev. Sci.* 22 (3), e12762. <https://doi.org/10.1111/desc.12762>.
- Ferjan Ramirez, N., Lytle, S.R., Kuhl, P.K., 2020. Parent coaching increases conversational turns and advances infant language development. *Proc. Natl. Acad. Sci.* 117 (7), 3484–3491.
- Fernald, A., Perfors, A., Marchman, V.A., 2006. Picking up speed in understanding: speech processing efficiency and vocabulary growth across the 2nd year. *Dev. Psychol.* 42 (1), 98–116. <https://doi.org/10.1037/0012-1649.42.1.98>.
- Fitch, W.T., Martins, M.D., 2014. Hierarchical processing in music, language, and action: lashley revisited. *Ann. N.Y. Acad. Sci.* 1316, 87–104. <https://doi.org/10.1111/nyas.12406>.
- Friederici, A.D., 2020. Hierarchy processing in human neurobiology: how specific is it? *Philos. Trans. R. Soc. B: Biol. Sci.* 375 (1789), 20180391. <https://doi.org/10.1098/rstb.2018.0391>.
- Fujiki, M., Brinton, B., Hart, C.H., Fitzgerald, A.H., 1999. Peer acceptance and Friendship in children with specific language impairment. *Top. Lang. Disord.* 19 (2), 34–48.
- Goldman, R., Fristoe, M., 2015. *WISC-III: Goldman-Fristoe 3 Test of Articulation*.
- Gordon, R.L., Shivers, C.M., Wieland, E.A., Kotz, S.A., Yoder, P.J., Devin McAuley, J., 2015. Musical rhythm discrimination explains individual differences in grammar skills in children. *Dev. Sci.* 18 (4), 635–644. <https://doi.org/10.1111/desc.12230>.
- Heard, M., Lee, Y.S., 2020. Shared neural resources of rhythm and syntax: an ALE meta-analysis. *Neuropsychologia* 137, 107284. <https://doi.org/10.1016/j.neuropsychologia.2019.107284>.
- Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8 (5), 393–402. <https://doi.org/10.1038/nrn2113>.
- Hubert-Dibon, G., Bru, M., Gras Le Guen, C., Launay, E., Roy, A., 2016. Health-related quality of life for children and adolescents with specific language impairment: a cohort study by a learning disabilities reference center. *Plos One* 11 (11), e0166541. <https://doi.org/10.1371/journal.pone.0166541>.
- Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., Kuhl, P.K., 2006. Infant speech perception activates broca's area: a developmental magnetoencephalography study. *Neuroreport* 17 (10), 957–962. <https://doi.org/10.1097/01.wnr.0000223387.51704.89>.
- Justice, L.M., Ahn, W.-Y., Logan, J.A.R., 2019. Identifying children with clinical language disorder: an application of machine-learning classification. *J. Learn. Disabil.* 52 (5), 351–365. <https://doi.org/10.1177/0022219419845070>.
- Kaufman, A.S., Kaufman, N.L., 2004. *Kaufman Assessment Battery for Children—Second Edition*. American Guidance Service, Circle Pines, MN.
- Kuhl, P.K., et al., 2005. Early speech perception and later language development: implications for the “critical period”. *Lang. Learn. Dev.* 1 (3–4), 237–264.
- Kuhl, P.K., 2010. Brain mechanisms in early language acquisition. *Neuron* 67 (5), 713–727. <https://doi.org/10.1016/j.neuron.2010.08.038>.
- Kuhl, P.K., 2021. Infant speech perception: integration of multimodal data leads to a new hypothesis—Sensorimotor mechanisms underlie learning. *Minnesota Symposia on Child Psychology*. Wiley, pp. 113–158.
- Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., Iverson, P., 2006. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* 9, F13–F21. <https://doi.org/10.1111/j.1467-7687.2006.00468.x>.
- Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., Nelson, T., 2008. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philos. Trans. R. Soc. B-Biol. Sci.* 363 (1493), 979–1000. <https://doi.org/10.1098/rstb.2007.2154>.
- Kuhl, P.K., Ramirez, R.R., Bosseler, A., Lin, J.-F.L., Imada, T., 2014. Infants' brain responses to speech suggest analysis by synthesis. *Proc. Natl. Acad. Sci.* <https://doi.org/10.1073/pnas.1410963111>.
- Kujala, T., Kallio, J., Tervaniemi, M., Naatanen, R., 2001. The mismatch negativity as an index of temporal processing in audition. *Clin. Neurophysiol.* 112 (9), 1712–1719. [https://doi.org/10.1016/s1388-2457\(01\)00625-3](https://doi.org/10.1016/s1388-2457(01)00625-3).
- Lancaster, H.S., Camarata, S., 2019. Reconceptualizing developmental language disorder as a spectrum disorder: issues and evidence. *Int. J. Lang. Commun. Disord.* 54, 79–94. <https://doi.org/10.1111/1460-6984.12433>.
- Law, J., Roy, P., 2008. Parental report of infant language skills: a review of the development and application of the communicative development inventories. *Child Adolescent Mental Health* 13, 198–206. <https://doi.org/10.1111/j.1475-3588.2008.00503.x>.
- Liégeois, F., Mayes, A., Morgan, A., 2014. Neural correlates of developmental speech and language disorders: evidence from neuroimaging. *Curr. Dev. Disord. Rep.* 1, 215–227. <https://doi.org/10.1007/s40474-014-0019-1>.
- Marchman, V.A., Fernald, A., 2008. Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Dev. Sci.* 11 (3), F9–F16. <https://doi.org/10.1111/j.1467-7687.2008.00671.x>.
- Martin, P., Murray, L.K., Darnell, D., Dorsey, S., 2018. Transdiagnostic treatment approaches for greater public health impact: implementing principles of evidence-based mental health interventions. *Clin. Psychol.* 25 (4), 1–13. <https://doi.org/10.1111/cpsp.12270>.
- Molfese, D.L., 2000. Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain Lang.* 72 (3), 238–245. <https://doi.org/10.1006/brln.2000.2287>.
- Molfese, D.L., Molfese, V.J., 1985. Electrophysiological indices of auditory discrimination in newborn infants: the bases for predicting later language development? *Infant Behav. Dev.* 8 (2), 197–211. [https://doi.org/10.1016/S0163-6383\(85\)80006-0](https://doi.org/10.1016/S0163-6383(85)80006-0).
- Molfese, D.L., Molfese, V.J., 1997. Discrimination of language skills at five years of age using event-related potentials recorded at birth. *Dev. Neuropsychol.* 13 (2), 135–156. <https://doi.org/10.1080/87565649709540674>.
- Naatanen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118 (12), 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>.
- National Academies of Sciences, Engineering Medicine, 2016. *Speech and Language Disorders in Children: Implications for the Social Security Administration's Supplemental Security Income Program*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/21872>.
- Newcomer, P.L., Hammill, D.D., 2008. *TOLD-P:4 : Test of Language Development*. Primary. PRO-ED, Austin TX.
- Newman, R., Ratner, N.B., Jusczyk, A.M., Jusczyk, P.W., Dow, K.A., 2006. Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Dev. Psychol.* 42 (4), 643–655. <https://doi.org/10.1037/0012-1649.42.4.643>.
- Newman, R.S., Rowe, M.L., Bernstein Ratner, N.A.N., 2016. Input and uptake at 7 months predicts toddler vocabulary: the role of child-directed speech and infant processing skills in language development. *J. Child Lang.* 43 (5), 1158–1173. <https://doi.org/10.1017/S0305000915000446>.
- Owen, A.J., Leonard, L.B., 2006. The production of finite and nonfinite complement clauses by children with specific language impairment and their typically developing peers. *J. Speech Lang. Hear. Res.* 49 (3), 548–571.
- Pallier, C., Devauchelle, A.-D., Dehaene, S., 2011. Cortical representation of the constituent structure of sentences. *Proc. Natl. Acad. Sci.* 108 (6), 2522–2527. <https://doi.org/10.1073/pnas.1018711108>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al., 2011. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Peter, B., Potter, N., Davis, J., et al., 2020. Toward a paradigm shift from deficit-based to proactive speech and language treatment: randomized pilot trial of the babble boot Camp in infants with classic galactosemia [version 5; Peer review: 2 approved, 1 approved with reservations]. *F1000Research* 8, 271. <https://doi.org/10.12688/f1000research.18062.5>.
- Preston, J.L., Felsenfeld, S., Frost, S.J., Mencl, W.E., 2012. Functional brain activation differences in school-age children with speech sound. *J. Speech Lang. Hear. Res.* 55 (August), 1068–1083. [https://doi.org/10.1044/1092-4388\(2011\)11-0056speech](https://doi.org/10.1044/1092-4388(2011)11-0056speech).
- Raghavan, R., Camarata, S., White, K., Barbaresi, W., Parish, S., Krahn, G., 2018. Population health in pediatric speech and language disorders: available data sources and a research agenda for the Field. *J. Speech Lang. Hear. Res.* 61 (5), 1279–1291. https://doi.org/10.1044/2018_JSLHR-L16-0459.
- Ramírez-Esparza, N., García-Sierra, A., Kuhl, P.K., 2014. Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Dev. Sci.* 17 (6), 880–891.
- Redmond, S.M., Thompson, H.L., Goldstein, S., 2011. Psycholinguistic profiling differentiates specific language impairment from typical development and from attention-Deficit/Hyperactivity disorder. *J. Speech Lang. Hear. Res.* 54 (1), 99–117.
- Rice, M., Wexler, K., 2001. *Rice Wexler Test of Early Grammatical Impairment*. Psychological Corporation, Hove.
- Rice, M., Taylor, C., Zubrick, S., 2008. Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *J. Speech Lang. Hear. Res.* 51 (2), 394–407.
- Sahin, N.T., Pinker, S., Cash, S.S., Schomer, D., Halgren, E., 2009. Sequential processing of lexical, grammatical, and phonological information within broca's Area. *Science* 326 (5951), 445–449. <https://doi.org/10.1126/science.1174481>.
- Sauer-Zavala, S., et al., 2017. Current definitions of “transdiagnostic” in treatment development: a search for consensus. *Behav. Ther.* 48 (1), 128–138.
- Segaert, K., Menenti, L., Weber, K., Petersson, K.M., Hagoort, P., 2011. Shared syntax in language production and language comprehension—an fMRI study. *Cereb. Cortex* 22 (7), 1662–1670. <https://doi.org/10.1093/cercor/bhr249>.
- Thal, D.J., O'Hanlon, L., Clemmons, M., Fralin, L., 1999. Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *J. Speech Lang. Hear. Res.* 42 (2), 482–496.
- Tkach, J.A., et al., 2011. Neural correlates of phonological processing in speech sound disorder: a functional magnetic resonance imaging study. *Brain Lang.* 119 (1), 42–49.
- Torgesen, J.K., Rashotte, C.A., Wagner, R.K., 2012. *TOWRE 2: Test of Word Reading Efficiency*. PRO-ED, Austin TX.

- Tsao, F.M., Liu, H.M., Kuhl, P.K., 2004. Speech perception in infancy predicts language development in the second year of life: a longitudinal study. *Child Dev.* 75 (4), 1067–1084. <https://doi.org/10.1111/j.1467-8624.2004.00726.x>.
- Tsao, F.-M., Liu, H.-M., Kuhl, P.K., 2006. Perception of native and non-native affricate- fricative contrasts: Cross-language tests on adults and infants. *J. Acoust. Soc. Am.* 120 (4), 2285–2294. <https://doi.org/10.1121/1.2338290>.
- Vabalas, A., Gowen, E., Poliakoff, E., Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. *Plos One* 14 (11), e0224365. <https://doi.org/10.1371/journal.pone.0224365>.
- Vissionon, K., Friederici, A.D., Brauer, J., Wu, C.-Y., 2017. Functional organization of the language network in three- and six-year-old children. *Neuropsychologia* 98, 24–33. <https://doi.org/10.1016/j.neuropsychologia.2016.08.014>.
- Werker, J.F., Curtin, S., 2005. PRIMIR: a developmental framework of infant speech processing. *Lang. Learn. Dev.* 1 (2), 197–234.
- Werker, J.F., Hensch, T.K., 2015. Critical periods in speech perception: New directions. *Annu. Rev. Psychol.* 66 (66), 173–196. <https://doi.org/10.1146/annurev-psych-010814-015104>.
- Werker, J.F., Tees, R.C., 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7, 49–63. [https://doi.org/10.1016/S0163-6383\(84\)80022-80023](https://doi.org/10.1016/S0163-6383(84)80022-80023).
- Xiao, Y., Brauer, J., Lauckner, M., Zhai, H., Jia, F., Margulies, D.S., Friederici, A.D., 2016. Development of the intrinsic language network in preschool children from ages 3 to 5 years. *Plos One* 11 (11), e0165802. <https://doi.org/10.1371/journal.pone.0165802>.
- Xie, Z., Reetzke, R., Chandrasekaran, B., 2019. Machine learning approaches to analyze speech-evoked neurophysiological responses. *J. Speech Lang. Hear. Res.* 62 (3), 587–601. <https://doi.org/10.1044/2018.JSLHR-S-ASTM-18-0244>.
- Zare, M., Rezvani, Z., Benasich, A.A., 2016. Automatic classification of 6-month-old infants at familial risk for language-based learning disorder using a support vector machine. *Clin. Neurophysiol.* 127 (7), 2695–2703. <https://doi.org/10.1016/j.clinph.2016.03.025>.
- Zhao, T.C., Kuhl, P.K., 2016. Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proc. Natl. Acad. Sci.* 113 (19), 5212–5217. <https://doi.org/10.1073/pnas.1603984113>.