

Neural oscillation coupling selectively predicts speech reception in young children with Autism Spectrum Disorder

Xiaoyue WANG^{1 5*}, Jaime DELGADO¹, Silvia MARCHESOTTI¹, Nada KOJOVIC², Holger Franz SPERDIN², Tonia A. RIHS³, Marie SCHAER^{2 4}, Anne-Lise GIRAUD^{1 4}

1 Auditory Language Group, Department of Basic Neuroscience, University of Geneva, Geneva, Switzerland

2 Autism Brain & Behavior Lab, Department of Psychiatry, University of Geneva, Geneva, Switzerland

3 Functional Brain Mapping Laboratory, Department of Basic Neuroscience, University of Geneva, Geneva, Switzerland

4 Co-last authors

5 Lead contact

*Correspondence: xiaoyue.wang@unige.ch

Highlights

Very young children diagnosed with autism spectrum disorder already present marked alterations of neural oscillatory activity in response to natural speech.

The hierarchical processing of phonemic- and syllabic-range information (theta/gamma coupling) is atypical in young ASD children.

Speech reception deficit in children with ASD is indexed by abnormal bottom-up (low-gamma) and top-down (low-beta) coordination.

Summary

Communication difficulties in autism spectrum disorder (ASD) involve a speech reception deficit, whose biological causes are not yet identified. This deficit could denote atypical neuronal ensemble activity, as reflected by neural oscillations. Atypical *cross-frequency oscillation coupling* in particular could disrupt the possibility to jointly track and predict dynamic acoustic stimuli, a dual process that is essential for speech comprehension. Whether

such oscillation anomalies can already be found in very young children with ASD, and with what specificity they relate to individual language reception capacity is unknown. In this study, neural activity was collected using EEG in 64 very young children with and without ASD (mean age 3) while they were exposed to naturalistic-continuous speech via an age-appropriate cartoon. EEG power typically associated with phrase-level chunking (delta, 1-3Hz), phonemic encoding (low-gamma, 25-35Hz) and top-down control (beta, 12-20Hz) was markedly reduced in ASD relative to typically developing (TD) children. Speech neural-tracking by delta and theta oscillations was also weaker in ASD than TD children. Critically, children with ASD exhibited slightly atypical theta/gamma coupling (PAC) involving a higher-than-normal gamma frequency, and markedly atypical beta/gamma PAC. Even though many oscillation features were atypical in our sample of 31 very young children with ASD, the beta/gamma coupling anomaly was the single best predictor of individual speech reception difficulties. These findings suggest that early interventions targeting the normalization of low-gamma and low-beta activity, might help young children with ASD to engage more in oral interactions.

Keywords Autism spectrum disorder (ASD) children, typically developing (TD) children, speech processing, low-frequency neural oscillations, neural tracking, phase-amplitude coupling (PAC)

Introduction

Although autism spectrum disorders (ASD) are primarily disorders of social interactions, language is seldom spared. Many people with ASD have severe language impairment, and even high functioning ASD individuals with excellent language skills retain difficulties understanding speech in noisy environments and/or when exposed to multiple speakers[1-4]. Individuals with ASD occasionally report that during childhood, speech was wholly or partly unintelligible[5, 6], and that speech reception was especially laborious with consonants, resulting in speech mostly sounding like a sequence of vowels[7]. Furthermore, even when they have an excellent language level, individuals with ASD often exhibit atypical-sounding elocution[8-12], which presumably further denotes speech reception anomalies[13, 14].

A large stream of recent studies shows that neural oscillations, i.e., the synchronous activity of neuronal populations, play a critical role in speech reception, mainly to parse the speech flow into meaningful linguistic units: phonemes, syllables, words, phrases, etc.[15-17]. While the 25-35Hz low-gamma oscillation rhythm is argued to work as a basic speech sampling rhythm enabling the encoding of phonemic-level acoustic features[18-20], the 4-7 Hz theta rhythm[21, 22] can flexibly track syllable boundaries[19, 23] hence playing a key role in speech intelligibility[24-28]. On the other hand, the slower 1-3 Hz delta rhythm has a more endogenous role in prosodic processing[29-31] and phrasal chunking[32, 33], while the 12-20Hz low-beta rhythm conveys top-down information presumably at an intermediate timescale between phonemes and syllables[23, 34]. Working in combination via their hierarchical coupling (the lower-frequency rhythm modulates the amplitude of higher-frequency activity), these different families of neural oscillations can underpin specific cognitive operations[35-37]. In particular, theta/gamma phase-amplitude coupling (PAC)[19, 38, 39] enables the hierarchical encoding of the phonemic structure within syllables, while beta/gamma PAC is associated with specific operations in predicting and planning speech[40-42].

In typically developing children, the ability to track the speech temporal structure develops very early[43-48]. The infant auditory system can already track the syllabic rhythm in nursery rhymes[44, 47] and native language sentences[46]. Even 1-8 days newborns can detect a consonant change on the sole basis of speech envelope cues[49]. Importantly, oscillation cross-frequency coupling (delta-theta/low-gamma), which orchestrates the encoding of

phonemic structure within syllables[19], was found in infants younger than 11 months[47]. Altogether, these findings suggest that the cortical tracking of speech rhythms is present at birth and likely contributes to phonological learning[50]. Its early disruption could hence be a possible cause of speech decoding impairment in children with ASD[21, 51], and participate in the altered development of their social skills.

Oscillation anomalies in response to speech have previously been reported in ASD, for example, decreased gamma responses to rapid spectrotemporal transitions associated with diphones[52-54], increased theta responses to repeated tones and syllables[55, 56], reduced beta responses to novel sounds[57, 58]. Young adults with ASD showed joint anomalies of neural theta (syllabic-level) and low-gamma (phonemic-level) activity, localizing predominantly to the auditory cortex. In this brain region, the relationship between theta and gamma neural activity was abnormal, suggesting that phonemic encoding was not temporally aligned with syllable tracking[51], a deficit that could deeply disrupt on-line speech reception[20]. Although the observed theta/gamma anomaly tightly correlated with the verbal scores of ASD participants[51], the study did not permit to conclude whether oscillation coupling anomalies are causally related to the speech reception deficit, in particular whether they are already present in very young children, when autistic symptoms usually become more evident, and at a development stage where typically developing children rapidly expand their speech expressive repertoire[59-61].

In this study, we used high-density electroencephalography (EEG) to compare oscillatory neural processing of age-appropriate naturalistic speech in 64 very young children (1.31- 5.56 years old) with and without ASD. Given the major developmental changes occurring in large-scale brain networks around this age in ASD[62], exploring speech reception as soon as the diagnosis is established is critical. If auditory cortical oscillations and their cross-frequency interactions are involved in language acquisition, we expect children diagnosed with ASD to already exhibit oscillation anomalies, notably in the delta-theta, low-beta, and low-gamma bands[23, 63-65]. We further expect these anomalies to specifically predict the language reception status (and not other developmental traits). The precise characterization of oscillation anomalies that have the potential to be causally involved in speech reception difficulties is an indispensable first step to envisage possible targeted interventions aiming at normalizing (boosting or down-regulating) oscillatory activity [66-68].

Results

Sixty-four children with (mean age: 3.09 ± 0.91) and without (mean age: 2.92 ± 1.30) ASD participated in an EEG experiment, with the consent of their legal representatives and in accordance with local Ethics regulations (Commission Cantonale d’Ethique de la Recherche, Geneva, Switzerland).

EEG and eye-gaze data were recorded while children watched an age-appropriate French cartoon[69-72]. Our aim was to capture speech-related neural activity in children with ASD compared to typically developing (TD) age matched peers while both groups were exposed to a naturalistic, attention grabbing social scene where one or more characters are speaking and interacting verbally. An acoustic characterization of the speech stimuli is shown in Figure S1. Visual attention to the cartoon was monitored by gaze data. Although the visual scene was not explored in the same way in each group [73], children with ASD watched the cartoon with attention (data with poor visual attention were excluded).

Speech-related oscillatory changes

To analyze EEG power and possible probe differences in neural activity in response to speech across groups, we computed the EEG power spectrum of speech compared to baseline in several frequency bands of interest, i.e., delta (1-3Hz), theta (4-8Hz), low-gamma (25-35Hz) and beta (12-20Hz). A between-group statistical comparison (Figure 1) showed reduced delta, beta, and low-gamma band oscillatory activity (mostly on mid-central clusters) in children with ASD. In contrast, theta oscillations were comparable to their TD peers.

Neural tracking of speech envelope using Multiple Linear Regression Model

We then explored the differences between ASD and TD in the neural tracking of the cartoon soundtrack’s phrase- and syllable-level modulations. We used a multiple linear regression (MLR) model with distributed lag between -300ms and 300ms with 50ms steps (Figure 2A) to reconstruct the stimulus envelopes from the neural signal in both the delta and the theta band. We then assessed the accuracy of speech envelope reconstruction by computing the correlation between reconstructed stimulus envelopes and actual envelopes in each group, before performing a statistical comparison. Speech envelope reconstruction was significantly

less accurate in ASD participants using the theta band signal (unpaired $t(62)=2.19$, $P = 0.04$, $\eta^2 = 0.07$, Figure 2B). The effect found in the theta band was most prominent in a specific cluster of 12 posterior-occipital electrodes (Figure 2C). The delta-band signal from all electrodes permitted to reconstruct the stimulus envelopes equally well in ASD and TD participants (unpaired $t(62)= 0.38$, $P = 0.71$, $\eta^2 = 0.002$, Figure 2B). However, we found reduced speech envelope reconstruction accuracy from the delta-band signal in ASD participants than controls, in a specific cluster of 7 parieto-occipital electrodes (Figure 2C). These results show that even though the overall theta power is unchanged in children with ASD (Figure 1), neural tracking of the speech syllabic structure by delta- and theta-range neural activity is altered.

Phase amplitude coupling

As speech is encoded hierarchically by different families of nested neural oscillations, we also analyzed phase-amplitude coupling (PAC) across modulating ($< 15\text{Hz}$) and modulated ($16\text{--}50\text{ Hz}$) frequencies using the KL-MI-Tort[74]. This approach was applied to those clusters showing significant between-group differences in the EEG power (all electrodes of clusters in the mid-central area, see Figure 1) and in the neural tracking analyses (theta-tracking cluster Figure 2C; delta-tracking cluster Figure S2A). Before computing PAC and assessing PAC changes, we checked that there were clear power spectrum peaks and troughs at each modulating frequency band of interest[75, 76]. Despite large inter-individual variability in peak/trough frequencies and oscillatory power (Figure S3), average neural activity confirmed significant power peaks in the low-frequency and the beta/gamma range in both groups (Figure 3A, 3C, Figure S2A), enabling us to compute comodulograms for speech and baseline EEG, and to compare them using cluster-based nonparametric statistics[77].

Although results from mid-central electrodes showed significant PAC clusters between the 3-8 Hz (phase) and the 22-38 Hz (amplitude) frequency ranges in both groups, there was only minimal overlap between ASD and TD (Figure 3B). While TD showed a unique and strong delta-theta/low-gamma PAC on mid-central (Figure 3B) and posterior-occipital electrodes (Figure 3D, Figure S2A), as typically observed in non-ASD adults and older children[19, 38, 39, 47], very young children with ASD exhibited wholly different patterns. The most significant difference was the extra presence of a consistent low-beta/low-gamma PAC in both mid-central and posterior-occipital electrodes, and an additional theta/gamma PAC in posterior occipital electrodes (Figure 3D, Figure S2A). In summary, compared with the TD group, children with ASD show atypical delta/theta-gamma coordination, and a robust low-

beta/low-gamma coupling anomaly suggesting abnormal endogenous control of speech reception.

Prediction of clinical variables from oscillatory neural features

Among all the oscillatory features that were found atypical in children with ASD, not all of them have the potential to specifically account for individual traits and, in particular, language abilities. We, therefore, assessed whether the group differences observed at the neurophysiological level could first predict the ASD severity, and more importantly, whether they could predict the speech reception scores obtained by the children in the MSEL, which cognitive scales measure overall developmental functioning. Conversely, we also sought to find out whether some of the oscillatory anomalies detected in power, tracking, and PAC analyses, could be more generally involved in several cognitive components (verbal production, visual processing, and fine motor skills).

Predicting ASD severity (ADOS severity level)

We first used nested LDA classification to determine which neurophysiological feature(s), i.e., EEG power, speech-tracking values, and PAC, best predicted the ASD symptom severity level, i.e., low, moderate, and high. Classification using all scalp electrodes showed that theta power (accuracy: 55.8%), low-gamma power (accuracy: 51.7%), delta-tracking (accuracy: 48.3%) were good predictors of autism severity (empirical chance level: 46.387%, Figure 4A).

In addition, low-beta/low-gamma PAC was a good predictor of autism severity in both clusters (accuracy 49.2% in the mid-central cluster and 51.7% in the posterior-occipital cluster, Figure 4B, Figure S2B, empirical chance level: 46.387%). These results demonstrate that theta power, gamma power, delta tracking, beta/gamma coupling contain critical information about autism severity.

Predicting speech reception level

We then used Lasso, with a nested cross-validation approach, to determine whether individual language skills could be predicted from oscillatory features and which features accounted most selectively for individual language development status. The analyses were run separately in each group. Although different in TD and ASD groups, band-specific EEG power accounted for none of the cognitive components of MSEL in any group, except language expression in TD children for the gamma power and language reception in children with ASD for the delta power (Figure 5A). Conversely, neural tracking, which was also

markedly different across groups, was predictive of several cognitive components in ASD (language reception and expression, fine motor skills) in the delta range, and of all but receptive language in the theta range (Figure 5B), suggesting the role of these oscillatory features in ASD goes beyond the sole domain of speech reception. In the TD group, neural tracking accounted for none of the scores, despite a trend for theta-tracking to predict both language reception and expression. In TD, the R^2 values were very low for all cognitive aspects in the delta range, and for motor and visual in the theta range (Figure 5B), showing that theta-tracking was more specifically involved in language development than delta, as expected from previous studies in adults[78, 79].

The most relevant feature for predicting language reception development was phase-amplitude coupling. As expected from studies in typical adults[38, 39] and from neurocomputational models[19], theta/gamma coupling selectively explained language reception in both central and posterior-occipital clusters of electrodes in our group of very young TD children (Figure 5C, Figure S5). In the ASD group, the (atypical) beta/gamma PAC selectively predicted language reception (Figure 5C). The post-hoc analysis of the dependency of language reception on beta/gamma PAC indicates that the stronger the anomaly, the worse speech reception (see Figure 5D). For displaying purposes, after confirming the performance of the algorithm, we used the whole data to run the hyperparameter optimization, and finally, we got the best estimator, then we retrained algorithm $y=f(x)$. We, therefore, obtained a fitted model $y = 17.45 \cdot f_p + 0.43 \cdot f_a - 5.71 \cdot MI - 142.19$, in which y represents language reception, f_p and f_a represent the frequency of phase and amplitude respectively, and MI refers to PAC value (Figure 5D). Overall, PAC was the most specific predictor of language reception in ASD and TD children: the presence of theta/gamma PAC predicts good speech reception in TD children confirming many previous observations[19, 38, 39, 47, 50], whereas the added presence of atypical beta/gamma PAC signals poor reception in ASD. Importantly, PAC features were much more sensitive than power and neural tracking to predict individual language reception scores.

Discussion

This study aimed to determine whether speech-related oscillatory anomalies in ASD are already present in early childhood, around the time of ASD diagnosis. Given the computational role neural oscillations play in chunking the syllable stream, encoding phonemic information, and predicting speech timing and linguistic content, we also sought to

determine whether and which of these anomalies most specifically account for speech reception difficulties in ASD. Establishing causality, or at least specific relevance of oscillation anomalies with respect to speech development, is critical, as early targeted neural interventions could subsequently be envisaged to normalize speech reception, as recently demonstrated in other neurodevelopmental language disorders[66, 80, 81]. Exploring EEG in 64 children between 1.31 and 5.56-years-old, we found marked anomalies of speech-induced cortical activity in the group with ASD, including decreased expression of delta, low-gamma, and beta frequencies. While theta power appeared as pronounced in ASD and TD children, delta-theta neural tracking was significantly reduced in ASD. Our most important results were observed in relation to oscillation cross-frequency coupling, which reflect the coordination computations at different timescales. As expected from previous studies in adults[19, 38, 39, 51], we clearly detected the classical theta/gamma coupling in very young (about 3-years-old) TD children and found this feature to specifically predict their individual language reception scores. This result represents an important finding as PAC is not a simple *marker* of speech reception ability but reflects a key computational component of speech processing, namely the hierarchical relationship between phonemes and syllables. This typical theta/low-gamma PAC was altered in children with ASD, appearing over a higher gamma range and a more occipital location. Critically, theta/gamma coupling was accompanied with a non-typical low-beta/low-gamma PAC in the ASD group, which, among all the abnormal oscillatory features reported here, was the only one from which we could specifically predict language reception scores in children with ASD: the stronger the beta/low-gamma coupling, the worse speech reception. This other key finding suggests that the speech processing computational scales are markedly different in ASD and that the endogenous-exogenous processing ratio is higher in children with ASD.

Low-gamma power predicts language expression in TD but not children with ASD

Our results showed decreased low-gamma activity in ASD relative to TD children, notably on mid-central electrodes, a scalp location that strongly captures auditory cortex activity[82, 83]. Previous studies already reported reduced gamma activity in response to pure tones[52, 53, 84], presumably denoting a basic functional anomaly of auditory cortex. Gamma activity usually reflects the excitation-inhibition balance[85] within brain circuits, which is a core parameter in neural development. Reduced low-gamma power in ASD could be hence seen as

a probe of atypical maturation of auditory neural circuitry. At the computational level, low-gamma activity is associated with phonemic encoding[18, 19] and its nesting within theta activity with the encoding of phonemic information within syllabic frames so that syllable-level representations can interface at the right time with other (higher) processing stages[86]. Here we found that low-gamma activity was the only feature that specifically predicted language expression in TD children, a logical finding as language expression is tightly related to the transformation of phonetic into articulatory features at the same timescale[87-91].

Abnormal theta-range speech tracking and theta/gamma coordination

Although we globally found a similar level of theta activity in both groups, theta-range *tracking* was reduced in children with ASD, meaning that theta activity although present did not typically follow the speech temporal structure. An equivalent level of global theta power in TD and ASD is consistent with previous studies[55, 56, 92]. However, that the speech envelope was less accurately reconstructed from the EEG in ASD than in TD children indicates that theta activity in ASD is more weakly engaged in syllable tracking[19, 21, 23, 93, 94]. This anomaly likely disrupts the alignment of neuronal excitability with syllabic onset, weakens the coordination with oscillations on other frequency bands, notably the low-gamma one, and ultimately hampers phonemic information encoding within syllables[18-20, 38, 39]. Accordingly, we found anomalies of the classical theta/gamma PAC in children with ASD: higher gamma frequencies and the occipital location are not compatible with a role in phonemic sampling within left auditory regions classically as reported in healthy individuals[17, 95-97]. These results however, are only in partial agreement with the more severe anomalies of theta/gamma coupling, a fully inverted coupling relationship, that were previously observed using simultaneous fMRI-EEG in adults with ASD[51]. Longitudinal studies are needed to determine whether theta/gamma coupling further deteriorates during childhood development and adolescence.

Theta-tracking of visual input

Abnormal theta tracking of speech syllable modulations might be associated or even follow from a stronger theta synchronization with the visual input. The posterior-occipital theta-tracking topography in children with ASD (Figure 2) could suggest that they track more

visual motion than auditory modulations, a functional anomaly that could possibly underpin atypical audio-visual integration[98-107]. This conjecture was partly confirmed by the direct comparison of visual and speech tracking (Figure S4), showing an imbalance in the visual vs. auditory tracking in the ASD group, with a predominant tracking of visual motion. In addition, the PAC frequency was higher than expected for speech processing, indicating that it is most likely local visual information, rather than auditory, that is modulated at theta rhythm. Overall, these results demonstrate a weak engagement of the auditory system in speech processing in very young children with ASD, and show that visual motion processing dominates when they watch a natural audio-visual scene involving moving and talking characters, as previously reported in adults/children[99-102].

Reduced delta-power but not delta-range speech tracking is a specific predictor of language reception in ASD

Our results also show reduced delta power and delta-range speech tracking in children with ASD. Delta-range activity signals phrase-level chunks, which do not necessarily have a physical/acoustic counterpart in the speech stimulus[79]. Delta activity is known to reflect a more endogenous processes than theta-range syllable tracking[64], which are argued to pertain to syntactic grouping[32, 63, 108, 109] or prosody processing[29, 30]. That reduced delta activity predicted speech reception in ASD might hence denote altered syntactic phrasal chunking and is compatible with previous observations suggesting both weaker linguistic[110] and intonation processing[111]. Although our results align well with these previous observations and hypotheses, delta-range speech tracking did not only predict speech reception scores but also expressive language, and fine motor skills. Logically, the more endogenous delta tracking deficit has a more global impact on the cognitive profile of children with ASD, and could hence also be a parameter that could possibly be adjusted using adapted neurostimulation methods.

Beta/gamma cross-frequency coupling: a speech reception singularity in young children with ASD

During continuous speech perception, top-down predictive mechanisms are also important, in particular in order to make sense of acoustic signals that might be unclear or ambiguous, or simply to follow the speaker's speech rate. Typically, predictive mechanisms are signaled by

the low-beta band[112, 113], which we also found to be weaker in ASD than TD children around the mid-central region. The beta frequency band is argued to mediate top-down information passing[42, 114, 115], allowing for precise temporal[65] but also content-specific predictions[42]. In speech processing, this frequency band could provide top-down integration constants that are intermediate between theta-syllabic and gamma-phonemic ranges[23, 34]. By occurring in alternation with bottom-up gamma phases[112], beta activity phases could stabilize representations in the face of the ever-changing acoustic input. Weaker beta activity in ASD might, therefore, reflect a reduced deployment of predictive mechanisms, altering both the ability to predict when acoustic signals can be expected and what content they carry. This interpretation is globally in line with alterations of phasic predictive learning in a mouse model of autism[116] and more generally with the hypothesis of impaired predictive coding in ASD[117, 118].

The key finding of this report is the atypical low-beta/low-gamma PAC found in children with ASD in conjunction with weaker beta power. A PAC anomaly involving the low-beta phase might follow from the reduced global beta power in ASD, possibly allowing low-gamma and gamma bursts to occur more strongly within beta troughs. Reduced beta gating could result in letting pass neural activity that is unrelated to the predicted speech structure, possibly leading to a feeling of being overwhelmed by unformatted acoustic inputs to which no linguistic value or meaning can be attributed. This interpretation aligns well with the auditory avoidance that is clinically described[110, 119, 120].

Cross-frequency oscillation features: a potential endophenotype for targeted interventions

As cortical oscillations arise from excitatory-inhibitory interactions within and across specific cortical laminae[121], auditory oscillation anomalies represent a plausible functional counterpart to structural disorganization and disruption of cortical inhibition previously shown in ASD[122]. A recent study directly relates neural oscillations with the expression of a number of genes, several of which are involved in ASD[123], e.g., LNX1, DGKI, KCNQ5, DCX, SHANK2, etc. Speech reception difficulties in ASD could, therefore, directly result from structural anomalies induced by mutations of genes[124-126] controlling neuronal interactions, notably at the synaptic level[125, 127].

Disruptions of synchronous neural activity in the cortex[128, 129] and other brain structures, such as the cerebellum and hippocampus[130-136], could lead to fragmented speech processing, abstraction difficulties in the auditory modality, and difficulties to map atypical auditory representations into appropriately timed articulatory sequences. Here, we found that among many neural oscillation anomalies, the most promising feature lies in cross-frequency coupling patterns, particularly the theta/gamma coupling that has an abnormal topography in ASD and the low-beta/low-gamma coupling that is wholly atypical. These two anomalies could be an ideal entry point for targeted brain stimulation interventions aiming at down-regulating the abnormal low-beta/low-gamma coupling, and re-localizing the typical auditory theta/low-gamma coupling to left auditory regions. The next step in this line of research will be to test whether a simple theta/gamma stimulation at the exact right frequencies (5Hz by 30 Hz) using, e.g., mild transcranial alternating current stimulation could indeed both disrupt low-beta/low-gamma and relocate theta/low-gamma activity to auditory regions. In combination with a close monitoring of behavior and EEG activity, such a trial would also allow us to firmly establish causality between auditory oscillatory activity and the speech reception ability in children with ASD.

Conclusion

Among all speech-related oscillatory alterations we report here in very young children diagnosed with ASD, cross-frequency features are those that predicted most closely individual language reception abilities. Cross-frequency coupling features hence appear as a promising language development endophenotype, bridging the gap from genetics to behavior, enabling accurate predictions of language development, and offering a precise entry point for specific interventions targeting the normalization of oscillatory functions.

Author Contributions

Data acquisition and clinical resources, M.S., N. K. and H.F.S.; Study design, X.W. and A-L.G.; Data analysis, XW, J.D. and S.M., T.A.R.; Writing-original draft, X.W., J.D., A-L.G.; Writing-review-editing, X.W., S.M., A-L.G. and M.S.; Supervision, A-L.G.

Acknowledgment

We thank Lucia Cantonas and Timothée Proix for their experts' technical support. This work

is supported by grants from the Swiss National Science Foundation (SNF 320030B_182855, 190084 and 51NF40_185897), and China Scholarship Council (Ph.D. grant to XW). The Cartool software (cartoolcommunity.unige.ch) has been programmed by Denis Brunet, from the Functional Brain Mapping Laboratory (FBMLab), Geneva, Switzerland, and is supported by the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne.

Declaration of interests

There is no conflict of interest.

Data and code availability

The unprocessed datasets for this manuscript are not publicly available yet, because analysis of these data is ongoing as part of a longitudinal study, and results are expected to be published in the future. When all data has been published, requests to access the datasets should be directed to Dr Marie Schaer, marie.schaer@unige.ch, the custom MATLAB analysis scripts will be made available upon request to the Lead Contact, Xiaoyue Wang (xiaoyue.wang@unige.ch)

Figure legends

Figure 1 Comparison of frequency-power in 31 children with ASD and 33 TD peers.

Children with ASD had reduced delta, beta, and low-gamma power and comparable theta power relative to their TD peers. Asterisks in topographies indicate group differences after cluster correction (cluster-based non-parametric permutation tests, cluster corrected $P = 0.05$).

Figure 2 Neural speech-tracking results.

(A) Schematics of the Multiple Linear Regression (MLR) model approach. The filtered EEG signal and stimulus envelope were entered into the MLR. For the cross-validation procedure, envelopes of speech and EEG signal of 9 folds with time-lag shifting were used to fit an MLR for each participant, which was then used to predict the envelope of the 10-th speech fold. The resulting model was tested by correlating the predicted envelope with the actual speech envelope in the test segment.

(B and C) Correlation coefficients between the reconstructed and real envelope, in all channels together (B) and each single-channel separately (C). Correlation coefficients were significantly reduced in ASD relative to TD in the delta and theta bands.

Error bars in B represent standard error, asterisks in C show group differences from non-parameter cluster-based permutation tests, $*P < 0.05$. See also Figure S4.

Figure 3 Phase-amplitude coupling.

(A and C) Speech-induced oscillatory power is characterized by a strong power peak in low-frequency bands (1-10 Hz) in both groups and a marked power trough in the low-beta band (10-15 Hz) in the ASD group over (A) central-electrodes selected from EEG power group differences and (C) posterior-occipital-electrodes selected from neural tracking group differences.

(B and D) Phase-amplitude comodulograms produced by statistically comparing modulation index (MI)-values in the speech and baseline periods over central-electrodes (B) and posterior-occipital-electrodes (D). Dotted lines represent significant differences in phase-amplitude coupling. For exact cluster locations, see topographies on A and C. For a quick appraisal of fp and fa ranges in each group, see the rightmost panel (fp : frequency of phase; fa : frequency of amplitude, nonparametric cluster-based statistics, cluster-corrected $P < 0.05$). See also Figure S2A.

Figure 4 Predicting ASD severity from EEG oscillatory activity.

(A) prediction accuracy of ASD symptom severity using the EEG power, neural tracking

based on all electrodes)

(B) prediction accuracy of ASD symptom severity using the phase-amplitude coupling (cluster-based)

The red line shows the chance level determined by an inverse binomial distribution. For the exact location of clusters, see topographies on top of panel B.

Figure 5 Predicting development level in young children from EEG oscillatory activity using a regularized linear model (Lasso).

(A and B) Low-gamma power significantly predicts Language expression in TD (A), delta-/theta-tracking significantly predict all tested cognitive components but language reception for theta tracking in ASD, and none in TD (B). (C) Theta/low-gamma PAC specifically predicted language reception in TD (C), whereas beta/low-gamma specifically predicted language reception in young children with ASD (C, left panel). R^2 values represent the proportion of the variance that is explained by the features for each target variable. (D) Language reception prediction from beta-gamma coupling ($r=-0.33$, $P = 0.04$).

Asterisks indicate the significant R^2 , $P<0.05$; β refers to low-beta, γ refers to low-gamma. For exact locations of clusters, see topographies on top of each panel in C. The missing bars indicate the R^2 is close to zero. See also Figure S5

Tables

Table 1. Participants' demographic information and group comparison of behavioral tests.

		Group					
		ASD (N=31, 6 females)		TD (N=33, 11 females)		t(df)	p
		Mean	SD	Mean	SD		
Age (in years)		3.09	0.91	2.92	1.30	0.62(62)	0.54
	Visual Reception	87.18	26.37	126.13	27.28	-5.53(56)	8.79E-7
Mullen Scales of Early Learning	Fine Motor	82.42	21.29	103.22	12.01	-4.54(56)	3.06E-5
	Expressive Language	58.52	31.69	104.77	24.08	-6.16(55)	9.18E-8
	Receptive Language	62.12	33.37	117.62	18.98	-7.71(56)	2.33E-10

- *Mullen Scales of Early Learning (assessment of cognitive and motor strengths and weakness in children).*
 - *Visual Reception & Fine Motor measure the child's nonverbal ability in order to estimate the overall developmental level.*
 - *Expressive language & Receptive Language measure the child's ability to process linguistic input and to use the language productively.*

STAR Methods

Inclusion procedure

Sixty-four children (mean age: 3.00 ± 1.12 , seventeen females) with and without ASD drawn from the Geneva Autism Cohort[137, 138] underwent a large battery of tests, including EEG recordings while they watches movies (French-cartoon for young children: TROTRO[69-72]). All children were recruited via specialized clinical centers or announcements in the community. The study was approved by the Ethics Committee of the Faculty of Medicine of the University of Geneva Hospital in accordance with the Declaration of Helsinki. A phone interview and a medical, developmental history questionnaire were completed before the initial visit of all participants; for the typically developing children, the following were considered exclusion criteria: any suspicion of altered development, a history of neurological or psychological disorder, and family history of ASD in first-degree relatives. Parents gave their informed consent prior to inclusion in the study. For all ASD children, the clinical diagnosis was formally corroborated using the Autism Diagnosis Observation Schedule- Generic (ADOS-G)[139] or the Autism Diagnosis Observation Schedule, second edition (ADOS-2)[140]. Data from 31 children with ASD (mean age: 3.09 ± 0.91) and 33 age-matched typically developing peers (mean age: 2.92 ± 1.30) were retained for data analyses (age difference: unpaired $t(62) = 0.62$, $P = 0.54$, see Table 1 and Supplementary Table1 for participant characteristics). Table 1 described the clinical characteristics of the ASD and the TD samples. Children with and without ASD were significantly impaired relative to their control peers across all used behavioral assessments, including language reception, language expression, visual reception, and fine motor.

Cognitive skills measure

Developmental functioning was assessed using the Mullen Scales of Early Learning (MSEL)[141], which comprise five subscales, namely gross motor, visual reception, fine motor, receptive language, and expressive language scales. The four latter scales are so-called “cognitive scales” and are used to derive an Early Learning Composite score as a measure of overall developmental functioning. The subscales of visual reception and fine motor skills measure nonverbal ability, while receptive language and expressive language measure the ability to process linguistic input and use the language productively. The child receives a “1” for correct response and “0” for incorrect response to age-adapted items. The subtest raw

scores were converted into age-adjusted normalized developmental quotient (DQ) scores, obtained by dividing the age-equivalent scores by the child's chronological age and multiplying the result by 100.

Symptom severity

Three symptom severity levels, i.e., low, moderate, high, were based on ADOS-G[139] or the ADOS-2[140] calibrated severity scores (Supplementary Table1), which varied from 3 to 10 distributed into three severity grades, 3-4 low-level; 5-7 moderate-level; 8-10 high-level.

Stimuli and Procedure

We explored cortical speech processing during a passive, naturalistic task with a relatively low cognitive demand suitable for young children. Participants watched the four French TROTRO[69-72] cartoon videos, each lasting about 2.5 minutes. The video presentation was controlled by Tobii Studio (Tobii® Technology, Sweden). Screen size was height \times width: 1200 pixels (29°38') \times 1920 pixels (45°53'), with a 60 Hz refresh rate. Participants were seated approximately 60 cm away from the screen. The cartoon soundtrack was delivered via loudspeakers at a sound level adjusted for each participant. The soundtrack sampling rate was 44.1 kHz. We removed background noise, e.g., birds' singing, music, from the original movie soundtrack using Audacity® v.2.2.1 editing software. We then extracted the stimulus envelopes using the absolute value of the analytic signal[142]. The speech envelope was down-sampled to 1000 Hz, low-pass filtered using a zero-phase fourth-order Butterworth filter set at 40 Hz (Figure S1A). An envelope spectral decomposition was performed using the Fast Fourier-transform (Figure S1B). We found dominant frequencies between 1 Hz and 7 Hz, with peaks at 1.17, 3.32, and 4.69 Hz, overlapping with the syllable rate range (4 to 6 syllables/sec, Figure S1C) as determined by averaging peaks within the 150ms minimum-peak-distance that was associated with the averaged French-syllable duration[143]. The spectrogram of the example sentence: "*je veux ta bougie rigolote en échange*" (I want your funny candle in exchange), shown in Figure 1D, was calculated using the Matlab function "spectrogram" (The MathWorks, Natick MA). Considering visual and auditory stimuli were embedded in natural, constantly changing, engaging videos, the visual motion envelope was also extracted from the stimulus video. Exemplary excerpts of visual motion are shown in Figure S1E,1F. To compute the average motion envelope across all pixels, we followed a method previously described in the literature[48, 144, 145]. Each video frame was converted

to greyscale, and the difference in luminance between two consecutive frames was computed. Then all pixels with a change larger than 10 (a threshold used to compensate for the noise of the video recording) were considered for computing the motion envelope as the average change value per video frame. These were done once for the length of the entire video, and the subsequent analysis retained only the portion of the video containing speech. On average, the change in luminance-derived motion envelope from one frame to the next was 0.88 units per frame (range 0–70.50, median=0.25). The motion envelope was then interpolated to the EEG sampling frequency of 1000 Hz. In all envelopes, time periods in which EEG data were excluded as a result of artifact rejection during preprocessing were zero-replaced. Finally, EEG data and visual motion envelopes corresponding to the speech parts of the soundtrack were aligned and available for the linear model analysis.

Eye-tracking acquisition and Pre-processing

Eye-gaze data were collected using a Tobii TX300 eye tracker (<https://www.tobii.com>), sampled at 300Hz. The cartoon frames subtended a visual angle of $26^{\circ}47' \times 45^{\circ}53'$ (height \times width). A five-point calibration procedure consisting of child-friendly animations was performed using an inbuilt program in the Tobii system. Upon verification, the calibration procedure was repeated if the eye-tracking device failed to detect the participant's gaze position accurately. The lighting conditions in the testing room were constant for all acquisitions. Participants younger than four years old sat on their parents' lap to make them feel comfortable and minimize head and body movements. All participants watched all four cartoon videos in the same order.

EEG acquisition and Pre-processing

EEG data were acquired using a Hydrocel Geodesic Sensor Net (HCGSN, Electrical Geodesics, USA) EEG system with 129 scalp electrodes at a 1000 Hz sampling rate. The reference electrode was located on the vertex (Cz), and a real-time band-pass filtered at 0-100 Hz was applied to the incoming signal. The first two cartoon videos were presented in a first block, and the last two cartoon videos in a second block, with a ~five-minute “dynamic image” task[146] in between cartoon videos and a 10-minutes break between blocks. At the end of the first block, the impedances were tested, and, if needed, the electrodes' conductance was adjusted by applying conductive gel or paste to keep impedances below 40 kOhm.

The signal pre-processing was conducted using the EEGlab v2019 toolbox within the

MATLAB environment[147] and Cartool (cartoolcommunity.unige.ch). First, the dataset was down-sampled to the montage of 110-channels by excluding cheek and neck electrodes that are often contaminated by muscle artifacts. Then, a zero-phase 4th order Butterworth bandpass filter between 0.1 and 70 Hz was applied to the EEG signals, as well as a notch filter at 50 Hz to remove power line interference. Each participant's data was then visually inspected to exclude periods contaminated by movement artifacts. An independent component analysis (ICA) was computed on the dataset to identify and remove components with eye blinks, saccades, electrical line noise, and heartbeat artifacts. The dimensionality of component was equivalent to the number of electrode arrays. Subsequently, a spherical spline interpolation was used to interpolate the channels contaminated by noise via the ICA-computed data. Finally, the Cartool spatial filter[148] was applied, and a common average reference was recalculated on the cleaned data. A trial was defined based on the beginning and end of each speech chunk in the cartoon, leading to a total of 50 trials with an average duration of 1.69 seconds. The detected artifacts periods were replaced by NaNs (not a number).

EEG power analysis

To identify the oscillatory responses to speech, a Morlet wavelet transformation was performed from 5 s before to 5 s after each speech chunk (making sure the epoch would not include any artifacts) period between 0.1 and 50 Hz with 5 cycles for Gaussian taper at each EEG electrode[149]. Then time-course between 1s pre-onset to 1s post-onset were selected (based on those speech chunks with the shortest duration) trial-by-trial, and the power was averaged across trials and normalized by decibel conversion (dB) over a -1000 ms to 0 ms baseline period, allowing for between-group comparisons. EEG power in the different frequency bands of interest were defined as the mean power value across 0-1000 ms post-stimulus onset. We compared ASD and TD groups in four frequency bands that are relevant for speech processing, i.e., theta (4-8Hz), delta (1-3Hz), beta (12-20Hz), and low-gamma (25-35Hz).

Speech envelope prediction from EEG power modulation: Multiple linear regression model (MLR) with distributed lags

To probe differences between ASD and TD in oscillatory speech tracking, a multiple linear

regression (MLR) model with a distributed lag between -300ms and 300ms with 50ms steps (Figure 2A) was used to reconstruct the stimulus envelopes from the neural responses. The EEG signals were lagged to compensate for possible differences in temporal alignment between the brain response and stimulus, a widely used method in the literature[150-153]. A multiple linear regression model was trained on the resulting set using a 10-fold cross-validation approach. First, all speech chunks were divided into ten consecutive segments. In each fold, one segment was left out for testing and the remaining segments were used for training. This process was repeated 10 times, ensuring that each segment was only used once for testing. On each fold, a multiple linear model was used to find a linear combination among the brain signals that best predicted the time course of the speech envelope. The resulting model was tested by correlating the predicted envelope with the actual speech envelope in the test segment. The final result, representing the oscillation tracking index, was obtained by averaging the correlation coefficient values in each fold.

We deployed the approach separately in the two frequency bands implicated in speech tracking, i.e., theta and delta, and computed the amplitude of specific bands to obtain the oscillation tracking index. The amplitude was the absolute value of the Hilbert transform in band-specific filtered EEG, i.e., 4-8 Hz for theta and 1-3 Hz for delta.

Phase-amplitude coupling

The time-courses from electrode clusters selected from the EEG power and neural tracking analyses were examined for phase-amplitude coupling changes (PAC). The first step was to obtain estimates of low-frequency phase (f_p) and high-frequency amplitude (f_a) by applying the Hilbert transform of each bandpass Butterworth filtered trial[154]. Considering the filters for extracting f_a must be wide enough to capture the center frequency \pm the modulating f_p to detect PAC[75, 76], we decided to use a variable bandwidth, defined as center frequency \pm the modulating center frequency, to improve PAC detection. The f_p bandwidth was kept narrow (center frequency \pm 1Hz) to extract sinusoidal waveforms. Furthermore, changes in the modulation power spectrum between speech and baseline periods was visually inspected in each participant to confirm that oscillation peaks/troughs were present at each modulating frequency-band of interest f_p . For instance, if the interested modulating frequency band was 4-8 Hz, we confirmed that participants presented a real peak in the power spectrum at 4-8 Hz. Next, the coupling between f_p and f_a was quantified using the Kullback–Leibler modulation index[74]. The KL-MI-Tort approach estimates PAC by quantifying the amount of deviation

in amplitude-phase distributions. This involves breaking f_p into 18 bins, and calculating the mean amplitude within each phase bin, normalized by the average value across all bins. Although the number of phase bins chosen is arbitrary, the specific number (9, 18, or 36) does not seem to influence PAC estimation[155]. The modulation index is calculated by comparing the amplitude-phase distribution (P) against the null hypothesis of a uniformly amplitude-phase distribution (Q).

$$MI = \frac{D(P,Q)}{\log(Nbin)} \quad (1)$$

Mathematically, this is computed using the Kullback–Leibler distance (D), related to Shannon’s entropy.

$$D(P, Q) = \sum_{ibin=1}^N P(ibin) \cdot \log\left(\frac{P(ibin)}{Q(ibin)}\right) \quad (2)$$

Using KL-MI-Tort, we calculated PAC between phases 2-15 Hz (in 1 Hz steps) and amplitudes 16-50 Hz (in 2 Hz steps) for the time-period 0–1000 ms following speech onset and a 1000 ms prestimulus baseline period. MI-values were calculated separately for each trial and averaged to obtain a single MI-value per amplitude and phase. To normalize MI values, this was repeated using surrogate data, created by shuffling trial and phase-carrying information (200 surrogates).

Predicting Clinical Variables from Oscillatory Features

In order to show the relationship between the brain activity and the neurophysiological variables, we tested whether autism severity was predicted by using only band-specific power per electrode (i.e., delta, theta, beta, and low-gamma power), or only neural tracking values per electrode, or only phase-amplitude coupling matrixes per cluster (i.e., maximum MI-value and corresponding phase-frequency, amplitude-frequency). For this, we trained a Linear Discriminant Analysis (LDA) classifier using a 10-fold nested cross-validation procedure, which separates the data in test and training sets. The training set was further separated using a 5-fold cross-validation approach for parameter search (Grid search). The LDA classifier was trained using a diagonal shared covariance matrix. The cross-validation process ensured that the training and testing data sets were not overlapping, avoiding misleading results due to overfitting. The input to the classifier was the band-specific EEG power, neural-tracking, and PAC for each participant, and the label to be predicted was symptom severity (i.e., low, moderate, and high). All empirical thresholds were obtained through cumulative binomial distribution[156].

Then, we probed the relationship between the neural and cognitive skills using a regularized linear model per frequency band (i.e., delta, theta, beta, and low-gamma power), per neural tracking per electrode, and per PAC matrix (i.e., maximum MI-value and corresponding phase-frequency, amplitude-frequency) to determine which critical oscillation (or combination of oscillations) was the best predictor of cognitive skills within group, e.g., the higher the tracking value, the higher the speech reception. The linear model was based on Lasso regression which requires finding the best hyper-parameters within high dimensional data[157]. We also used a 10-fold nested cross-validation approach to improve model selection. The training set was further separated using a 5-fold cross-validation approach for parameter search (Grid search). The results were presented as averaged R^2 values, which indicate the prediction power of a given feature, i.e., higher R^2 's signal higher prediction accuracy.

Statistical analysis

Between-group statistical comparisons of band-specific EEG power and neural tracking were done using cluster-based nonparametric permutation tests with Monte Carlo randomization[77] using the FieldTrip toolbox[158]; <http://fieldtriptoolbox.org>). For details, between-groups differences (ASD vs. TD) were assessed with unpaired t -statistics. First, clusters of significant group differences were obtained by considering at least two adjacent electrodes whose t -value exceeded a 5% significance threshold (ASD vs. TD, unpaired t -test, uncorrected for multiple comparisons). The maximum t -value within each cluster was carried forward. Next, a null distribution was obtained by randomizing the group label 1,000 times and calculating the largest cluster-level t -value for each permutation. The maximum t -value within each original cluster was then compared against this null distribution, with values exceeding a threshold of $p < 0.05$ deemed significant. Besides, between-group differences of neural tracking were tested by means of unpaired t -test for all-electrodes analyses.

To assess changes in the comodulograms of PAC between the speech and baseline periods, then we repeated the nonparametric cluster-based statistics[77], except the first step that uncorrected dependent-samples t -test was performed (speech vs. baseline), and all MI-values exceeding a 5% significance threshold were grouped into clusters.

Supplementary information (in separate file)

References

1. Dunlop, W.A., Enticott, P.G., and Rajan, R. (2016). Speech Discrimination Difficulties in High-Functioning Autism Spectrum Disorder Are Likely Independent of Auditory Hypersensitivity. *Front Hum Neurosci* *10*, 401.
2. Alcantara, J.I., Weisblatt, E.J., Moore, B.C., and Bolton, P.F. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *J Child Psychol Psychiatry* *45*, 1107-1114.
3. Schelinski, S., and von Kriegstein, K. (2020). Brief Report: Speech-in-Noise Recognition and the Relation to Vocal Pitch Perception in Adults with Autism Spectrum Disorder and Typical Development. *Journal of autism and developmental disorders* *50*, 356-363.
4. Beker, S., Foxe, J.J., and Molholm, S. (2018). Ripe for solution: Delayed development of multisensory processing in autism and its remediation. *Neurosci Biobehav Rev* *84*, 182-192.
5. Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., Murphy, S., and Almirall, D. (2014). Communication interventions for minimally verbal children with autism: a sequential multiple assignment randomized trial. *J Am Acad Child Adolesc Psychiatry* *53*, 635-646.
6. Grandin, T. (2011). *The way I see it: A personal look at autism & Asperger's*, (Future Horizons).
7. McKeever, L., Cleland, J., and Delafield-Butt, J. (2019). Aetiology of speech sound errors in autism. In *Speech production and perception: Learning and memory*, S. Fuchs, J. Cleland and A. Rochet-Capellan, eds. (Bern, Switzerland: Peter Lang).
8. Grandin, T. (1995). How People with Autism Think. In *Learning and Cognition in Autism*, E. Schopler and G.B. Mesibov, eds. (Boston, MA: Springer US), pp. 137-156.
9. Grandin, T. (2010). *The world needs all kinds of minds*.
10. Grandin, T. (2008). *Thinking in pictures, expanded edition: My life with autism*, (Vintage).
11. Nadig, A., and Shaw, H. (2012). Acoustic and Perceptual Measurement of Expressive Prosody in High-Functioning Autism: Increased Pitch Range and What it Means to Listeners. *Journal of autism and developmental disorders* *42*, 499-511.
12. Shriberg, L.D., Paul, R., McSweeney, J.L., Klin, A.M., Cohen, D.J., and Volkmar, F.R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *J Speech Lang Hear Res* *44*, 1097-1115.
13. Akahane-Yamada, R., Tohkura, Y., Bradlow, A.R., and Pisoni, D.B. (1996). Does training in speech perception modify speech production? In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96, Volume 2*. pp. 606-609 vol.602.
14. Paul, R., Augustyn, A., Klin, A., and Volkmar, F.R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of autism and developmental disorders* *35*, 205-220.
15. Arnal, L.H., Poeppel, D., and Giraud, A.L. (2015). Temporal coding in the auditory cortex. *Handb Clin Neurol* *129*, 85-98.
16. Greenberg, S., Carvey, H., Hitchcock, L., and Chang, S. (2003). Temporal properties of spontaneous speech—a syllable-centric perspective. *Journal of Phonetics* *31*, 465-485.
17. Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun* *41*, 245-255.
18. Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., and Giraud, A.L. (2011). Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron* *72*, 1080-1090.
19. Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., and Giraud, A.L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *Elife* *4*, e06213.
20. Hovsepyan, S., Olasagasti, I., and Giraud, A.L. (2020). Combining predictive coding and neural oscillations enables online syllable recognition in natural speech. *Nat Commun* *11*, 3117.
21. Giraud, A.L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci* *15*, 511-517.
22. Ding, N., Patel, A., Chen, L., Butler, H., Luo, C., and Poeppel, D. (2016). Temporal Modulations Reveal Distinct Rhythmic Properties of Speech and Music. *BioRxiv*, 059683.
23. Pefkou, M., Arnal, L.H., Fontolan, L., and Giraud, A.L. (2017). theta-Band and beta-Band Neural Activity Reflects Independent Syllable Tracking and Comprehension of Time-Compressed Speech. *The Journal of neuroscience : the official journal of the Society for Neuroscience* *37*, 7930-7938.
24. Zoefel, B., and VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage* *124*, 16-23.
25. Zoefel, B. (2018). Speech Entrainment: Rhythmic Predictions Carried by Neural Oscillations. *Curr Biol* *28*, R1102-R1104.
26. Peelle, J.E., Gross, J., and Davis, M.H. (2013). Phase-locked responses to speech in human auditory

- cortex are enhanced during comprehension. *Cereb Cortex* 23, 1378-1387.
27. Di Liberto, G.M., O'Sullivan, J.A., and Lalor, E.C. (2015). Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing. *Curr Biol* 25, 2457-2465.
 28. Blank, H., and Davis, M.H. (2016). Prediction Errors but Not Sharpened Signals Simulate Multivoxel fMRI Patterns during Speech Perception. *PLoS Biol* 14, e1002577.
 29. Teoh, E.S., Cappelloni, M.S., and Lalor, E.C. (2019). Prosodic pitch processing is represented in delta-band EEG and is dissociable from the cortical tracking of other acoustic and phonetic features. *Eur J Neurosci* 50, 3831-3842.
 30. Ghitza, O. (2017). Acoustic-driven delta rhythms as prosodic markers. *Lang Cogn Neurosci* 32, 545-561.
 31. Bourguignon, M., De Tieghe, X., de Beeck, M.O., Ligot, N., Paquier, P., Van Bogaert, P., Goldman, S., Hari, R., and Jousmaki, V. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Hum Brain Mapp* 34, 314-326.
 32. Ding, N., Chatterjee, M., and Simon, J.Z. (2014). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage* 88, 41-46.
 33. Rimmele, J.M., Poeppel, D., and Ghitza, O. (2020). Acoustically driven cortical delta oscillations underpin perceptual chunking. *bioRxiv*.
 34. Giraud, A.L. (2020). Oscillations for all A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 1-8.
 35. Hyafil, A., Giraud, A.L., Fontolan, L., and Gutkin, B. (2015). Neural Cross-Frequency Coupling: Connecting Architectures, Mechanisms, and Functions. *Trends in neurosciences* 38, 725-740.
 36. Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., Berger, M.S., Barbaro, N.M., and Knight, R.T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313, 1626-1628.
 37. Canolty, R.T., and Knight, R.T. (2010). The functional role of cross-frequency coupling. *Trends Cogn Sci* 14, 506-515.
 38. Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., and Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol* 11, e1001752.
 39. Lizarazu, M., Lallier, M., and Molinaro, N. (2019). Phase-amplitude coupling between theta and gamma oscillations adapts to speech rate. *Ann N Y Acad Sci*.
 40. Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *Eur J Neurosci* 48, 2609-2621.
 41. Livezey, J.A., Bouchard, K.E., and Chang, E.F. (2019). Deep learning as a tool for neural data analysis: Speech classification and cross-frequency coupling in human sensorimotor cortex. *Plos Comput Biol* 15.
 42. Chao, Z.C., Takaura, K., Wang, L., Fujii, N., and Dehaene, S. (2018). Large-Scale Cortical Networks for Hierarchical Prediction and Prediction Error in the Primate Brain. *Neuron* 100, 1252-1266 e1253.
 43. Cabrera, L., and Gervain, J. (2020). Speech perception at birth: The brain encodes fast and slow temporal information. *Science Advances* 6, eaba7830.
 44. Leong, V., Byrne, E., Clackson, K., Harte, N., Lam, S., Barbaro, K.d., and Wass, S. (2017). Infants' neural oscillatory processing of theta-rate speech patterns exceeds adults'. *bioRxiv*, 108852.
 45. Kalashnikova, M., Peter, V., Di Liberto, G.M., Lalor, E.C., and Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Sci Rep* 8, 13745.
 46. Ortiz Barajas, M.C., Guevara Erra, R., and Gervain, J. (2021). The origins and development of speech envelope tracking during the first months of life. *Developmental Cognitive Neuroscience*.
 47. Attaheri, A., Choidealbha, Á.N., Di Liberto, G.M., Rocha, S., Brusini, P., Mead, N., Olawole-Scott, H., Boutris, P., Gibbon, S., Williams, I., et al. (2020). Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants.
 48. Jessen, S., Fiedler, L., Münte, T.F., and Obleser, J. (2019). Quantifying the individual auditory and visual brain response in 7- month-old infants watching a brief cartoon movie. *NeuroImage*.
 49. Cabrera, L., and Gervain, J. (2019). How infants use the envelope of the speech signal to perceive phonetic contrasts at birth. *PsyArXiv*.
 50. Goswami, U. (2019). Speech rhythm and language acquisition: an amplitude modulation phase hierarchy perspective. *Ann N Y Acad Sci* 1453, 67-78.
 51. Jochaut, D., Lehongre, K., Saitovitch, A., Devauchelle, A.D., Olasagasti, I., Chabane, N., Zilbovicius, M., and Giraud, A.L. (2015). Atypical coordination of cortical oscillations in response to speech in autism. *Front Hum Neurosci* 9, 171.
 52. Rojas, D.C., Maharajh, K., Teale, P., and Rogers, S.J. (2008). Reduced neural synchronization of gamma-band MEG oscillations in first-degree relatives of children with autism. *BMC Psychiatry* 8, 66.
 53. Gandal, M.J., Edgar, J.C., Ehrlichman, R.S., Mehta, M., Roberts, T.P.L., and Siegel, S.J. (2010).

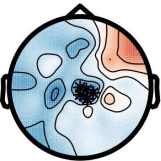
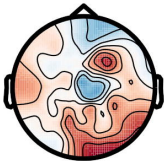
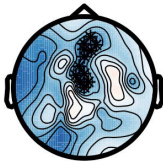
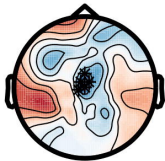
- Validating gamma Oscillations and Delayed Auditory Responses as Translational Biomarkers of Autism. *Biol Psychiat* 68, 1100-1106.
54. Ortiz-Mantilla, S., Cantiani, C., Shafer, V.L., and Benasich, A.A. (2019). Minimally-verbal children with autism show deficits in theta and gamma oscillations during processing of semantically-related visual information. *Sci Rep-Uk* 9.
55. Zhang, J., Meng, Y., Tong, X., Yuan, Z., Wu, C., and Jeong, S.L. (2018). Exploring the neural correlates of lexical stress perception in english among Chinese-English bilingual children with autism spectrum disorder: An ERP study. *Neurosci Lett* 666, 158-164.
56. Wang, X., Wang, S., Fan, Y., Huang, D., and Zhang, Y. (2017). Speech-specific categorical perception deficit in autism: An Event-Related Potential study of lexical tone processing in Mandarin-speaking children. *Sci Rep* 7, 43254.
57. Hames, E.C., Murphy, B., Rajmohan, R., Anderson, R.C., Baker, M., Zupancic, S., O'Boyle, M., and Richman, D. (2016). Visual, Auditory, and Cross Modal Sensory Processing in Adults with Autism: An EEG Power and BOLD fMRI Investigation. *Front Hum Neurosci* 10, 167.
58. Mamashli, F., Khan, S., Bharadwaj, H., Michmizos, K., Ganesan, S., Garel, K.L.A., Ali Hashmi, J., Herbert, M.R., Hämäläinen, M., and Kenet, T. (2017). Auditory processing in noise is associated with complex patterns of disrupted functional connectivity in autism spectrum disorder. *Autism Research* 10, 631-647.
59. Hoff, E. (2013). *Language development*, (Cengage Learning).
60. Press, R. (2015). *Developmental milestones of young children*, (Redleaf Press).
61. Whitehead, M.R. (2010). *Language & literacy in the early years 0-7*, (Sage).
62. Nomi, J.S., and Uddin, L.Q. (2015). Developmental changes in large-scale network connectivity in autism. *Neuroimage Clin* 7, 732-741.
63. Meyer, L., Henry, M.J., Gaston, P., Schmuck, N., and Friederici, A.D. (2017). Linguistic Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. *Cereb Cortex* 27, 4293-4302.
64. Molinaro, N., and Lizarazu, M. (2018). Delta (but not theta)-band cortical entrainment involves speech-specific processing. *Eur J Neurosci* 48, 2642-2650.
65. Fujioka, T., Trainor, L.J., Large, E.W., and Ross, B. (2012). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *Journal of Neuroscience* 32, 1791-1802.
66. Marchesotti, S., Nicolle, J., Merlet, I., Arnal, L.H., Donoghue, J.P., and Giraud, A.L. (2020). Selective enhancement of low-gamma activity by tACS improves phonemic processing and reading accuracy in dyslexia. *PLoS Biol* 18, e3000833.
67. Kayarian, F.B., Jannati, A., Rotenberg, A., and Santarnecchi, E. (2020). Targeting Gamma-Related Pathophysiology in Autism Spectrum Disorder Using Transcranial Electrical Stimulation: Opportunities and Challenges. *Autism Res*.
68. Jones, K.T., Johnson, E.L., Tauxe, Z.S., and Rojas, D.C. (2020). Modulation of auditory gamma-band responses using transcranial electrical stimulation. *J Neurophysiol* 123, 2504-2514.
69. Lezoray, S. (2013). Trotro es tres amoureux. In *Trotro*. (Storimages), p. 2.5 Minutes.
70. Lezoray, S. (2013). Trotro et la boite a secrets. In *Trotro*. (Storimages), p. 2.5 Minutes.
71. Lezoray, S. (2013). Trotro part en vacance. In *Trotro*. (Storimages), p. 2.5 Minutes.
72. Lezoray, S. (2013). Trotro l'anniversaire de nana. In *Trotro*. (Storimages), p. 2.5 Minutes.
73. Kojovic, N., Franchini, M., Sperdin, H.F., Sandini, C., Jan, R.K., Zöllner, D., and Schaer, M. (2020). Unraveling the Developmental Dynamic of Visual Exploration of Social Interactions in Autism.
74. Tort, A.B., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *J Neurophysiol* 104, 1195-1210.
75. Dvorak, D., and Fenton, A.A. (2014). Toward a proper estimation of phase-amplitude coupling in neural oscillations. *J Neurosci Methods* 225, 42-56.
76. Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., Singer, W., and Vicente, R. (2015). Untangling cross-frequency coupling in neuroscience. *Curr Opin Neurobiol* 31, 51-61.
77. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164, 177-190.
78. Vander Ghinst, M., Bourguignon, M., Niesen, M., Wens, V., Hassid, S., Choufani, G., Jousmaki, V., Hari, R., Goldman, S., and De Tieghe, X. (2019). Cortical Tracking of Speech-in-Noise Develops from Childhood to Adulthood. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 39, 2938-2950.
79. Boucher, V.J., Gilbert, A.C., and Jemel, B. (2019). The Role of Low-frequency Neural Oscillations in Speech Processing: Revisiting Delta Entrainment. *J Cogn Neurosci* 31, 1205-1215.
80. Meyer, L., Lakatos, P., and He, Y. (2021). Language Dysfunction in Schizophrenia: Assessing Neural Tracking to Characterize the Underlying Disorder(s)? *Front Neurosci* 15, 640502.
81. Ladanyi, E., Persici, V., Fiveash, A., Tillmann, B., and Gordon, R.L. (2020). Is atypical rhythm a risk

- factor for developmental speech and language disorders? *Wiley Interdiscip Rev Cogn Sci*, e1528.
82. Stropahl, M., Bauer, A.K.R., Debener, S., and Bleichner, M.G. (2018). Source-Modeling Auditory Processes of EEG Data Using EEGLAB and Brainstorm. *Front Neurosci-Switz* 12.
 83. Steinschneider, M., Liégeois-Chauvel, C., and Brugge, J.F. (2011). Auditory Evoked Potentials and Their Utility in the Assessment of Complex Sound Processing. In *The Auditory Cortex*, J.A. Winer and C.E. Schreiner, eds. (Boston, MA: Springer US), pp. 535-559.
 84. Edgar, J.C., Khan, S.Y., Blaskey, L., Chow, V.Y., Rey, M., Gaetz, W., Cannon, K.M., Monroe, J.F., Cornew, L., Qasmieh, S., et al. (2015). Neuromagnetic oscillations predict evoked-response latency delays and core language deficits in autism spectrum disorders. *Journal of autism and developmental disorders* 45, 395-405.
 85. Buzsaki, G., and Wang, X.J. (2012). Mechanisms of gamma oscillations. *Annu Rev Neurosci* 35, 203-225.
 86. Ghitza, O. (2011). Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Front Psychol* 2, 130.
 87. Proix, T., Saa, J.D., Christen, A., Martin, S., Pasley, B.N., Knight, R.T., Tian, X., Poeppel, D., Doyle, W.K., Devinsky, O., et al. (2021). Imagined speech can be decoded from low- and cross-frequency features in perceptual space. *bioRxiv*, 2021.2001.2026.428315.
 88. Cheung, C., Hamiton, L.S., Johnson, K., and Chang, E.F. (2016). The auditory representation of speech sounds in human motor cortex. *Elife* 5.
 89. Tang, C., Hamilton, L.S., and Chang, E.F. (2017). Intonational speech prosody encoding in the human auditory cortex. *Science* 357, 797-801.
 90. Yi, H.G., Leonard, M.K., and Chang, E.F. (2019). The Encoding of Speech Sounds in the Superior Temporal Gyrus. *Neuron* 102, 1096-1110.
 91. Chartier, J., Anumanchipalli, G.K., Johnson, K., and Chang, E.F. (2018). Encoding of Articulatory Kinematic Trajectories in Human Speech Sensorimotor Cortex. *Neuron* 98, 1042-1054 e1044.
 92. Ahn, S., Cho, H., Kwon, M., Kim, K., Kwon, H., Kim, B.S., Chang, W.S., Chang, J.W., and Jun, S.C. (2018). Interbrain phase synchronization during turn-taking verbal interaction—a hyperscanning study using simultaneous EEG/MEG. *Human brain mapping* 39, 171-188.
 93. Ghitza, O., and Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica* 66, 113-126.
 94. Luo, H., and Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54, 1001-1010.
 95. Giraud, A.L., Kleinschmidt, A., Poeppel, D., Lund, T.E., Frackowiak, R.S., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56, 1127-1134.
 96. Morillon, B., Lehongre, K., Frackowiak, R.S., Ducorps, A., Kleinschmidt, A., Poeppel, D., and Giraud, A.L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proc Natl Acad Sci U S A* 107, 18688-18693.
 97. Telkemeyer, S., Rossi, S., Koch, S.P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29, 14726-14733.
 98. Simon, D.M., and Wallace, M.T. (2016). Dysfunction of sensory oscillations in Autism Spectrum Disorder. *Neurosci Biobehav Rev* 68, 848-861.
 99. Feldman, J.I., Dunham, K., Cassidy, M., Wallace, M.T., Liu, Y., and Woynaroski, T.G. (2018). Audiovisual multisensory integration in individuals with autism spectrum disorder: A systematic review and meta-analysis. *Neurosci Biobehav Rev* 95, 220-234.
 100. Stevenson, R.A., Segers, M., Ferber, S., Barense, M.D., Camarata, S., and Wallace, M.T. (2016). Keeping time in the brain: Autism spectrum disorder and audiovisual temporal processing. *Autism Res* 9, 720-738.
 101. Stevenson, R.A., Siemann, J.K., Schneider, B.C., Eberly, H.E., Woynaroski, T.G., Camarata, S.M., and Wallace, M.T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 34, 691-697.
 102. Foss-Feig, J.H., Kwakye, L.D., Cascio, C.J., Burnette, C.P., Kadivar, H., Stone, W.L., and Wallace, M.T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Exp Brain Res* 203, 381-389.
 103. Sumby, W.H., and Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *Journal of the Acoustical Society of America* 26, 212-215.
 104. Simon, D.M., and Wallace, M.T. (2017). Rhythmic Modulation of Entrained Auditory Oscillations by Visual Inputs. *Brain Topogr* 30, 565-578.

105. Demiralp, T., Bayraktaroglu, Z., Lenz, D., Junge, S., Busch, N.A., Maess, B., Ergen, M., and Herrmann, C.S. (2007). Gamma amplitudes are coupled to theta phase in human EEG during visual perception. *Int J Psychophysiol* *64*, 24-30.
106. Luo, H., Liu, Z., and Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biol* *8*, e1000445.
107. Micheli, C., Schepers, I.M., Ozker, M., Yoshor, D., Beauchamp, M.S., and Rieger, J.W. (2018). Electroencephalography reveals continuous auditory and visual speech tracking in temporal and occipital cortex. *Eur J Neurosci*.
108. Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nat Neurosci* *19*, 158-164.
109. Keitel, A., Gross, J., and Kayser, C. (2017). Speech tracking in auditory and motor regions reflects distinct linguistic features. *bioRxiv*, 195941.
110. Haesen, B., Boets, B., and Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Research in Autism Spectrum Disorders* *5*, 701-714.
111. Benitez-Burraco, A., and Murphy, E. (2016). The Oscillopathic Nature of Language Deficits in Autism: From Genes to Language Evolution. *Front Hum Neurosci* *10*, 120.
112. Fontolan, L., Morillon, B., Liegeois-Chauvel, C., and Giraud, A.L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nat Commun* *5*, 4694.
113. Park, H., Ince, R.A., Schyns, P.G., Thut, G., and Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Curr Biol* *25*, 1649-1653.
114. Wang, X.J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev* *90*, 1195-1268.
115. Bastos, A.M., Lundqvist, M., Waite, A.S., Kopell, N., and Miller, E.K. (2020). Layer and rhythm specificity for predictive routing. *Proc Natl Acad Sci U S A* *117*, 31459-31469.
116. Kosaki, Y., and Watanabe, S. (2016). Impaired Pavlovian predictive learning between temporally phasic but not static events in autism-model strain mice. *Neurobiol Learn Mem* *134 Pt B*, 304-316.
117. Sinha, P., Kjelgaard, M.M., Gandhi, T.K., Tsourides, K., Cardinaux, A.L., Pantazis, D., Diamond, S.P., and Held, R.M. (2014). Autism as a disorder of prediction. *Proc Natl Acad Sci U S A* *111*, 15220-15225.
118. Lieder, I., Adam, V., Frenkel, O., Jaffe-Dax, S., Sahani, M., and Ahissar, M. (2019). Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. *Nat Neurosci* *22*, 256-264.
119. Kamita, M.K., Silva, L.A.F., Magliaro, F.C.L., Fernandes, F.D., and Matas, C.G. (2021). Auditory Event Related Potentials in children with autism spectrum disorder. *Int J Pediatr Otorhinolaryngol* *148*, 110826.
120. Top, D.N., Jr., Luke, S.G., Stephenson, K.G., and South, M. (2018). Psychophysiological Arousal and Auditory Sensitivity in a Cross-Clinical Sample of Autistic and Non-autistic Anxious Adults. *Front Psychiatry* *9*, 783.
121. Cannon, J., McCarthy, M.M., Lee, S., Lee, J., Borgers, C., Whittington, M.A., and Kopell, N. (2014). Neurosystems: brain rhythms and cognitive processing. *Eur J Neurosci* *39*, 705-719.
122. Rojas, D.C., Singel, D., Steinmetz, S., Hepburn, S., and Brown, M.S. (2014). Decreased left perisylvian GABA concentration in children with autism and unaffected siblings. *NeuroImage* *86*, 28-34.
123. Berto, S., Fontenot, M.R., Seger, S., Ayhan, F., Caglayan, E., Kulkarni, A., Douglas, C., Tamminga, C.A., Lega, B.C., and Konopka, G. (2021). Gene-expression correlates of the oscillatory signatures supporting human episodic memory encoding. *Nat Neurosci*.
124. Nguyen, T.A., Lehr, A.W., and Roche, K.W. (2020). Neuroligins and Neurodevelopmental Disorders: X-Linked Genetics. *Front Synaptic Neurosci* *12*, 33.
125. Caubit, X., Gubellini, P., Andrieux, J., Roubertoux, P.L., Metwaly, M., Jacq, B., Fatmi, A., Had-Aissouni, L., Kwan, K.Y., Salin, P., et al. (2016). TSHZ3 deletion causes an autism syndrome and defects in cortical projection neurons. *Nat Genet* *48*, 1359-1369.
126. Boeckx, C., and Benitez-Burraco, A. (2014). The shape of the human language-ready brain. *Front Psychol* *5*, 282.
127. Peca, J., Feliciano, C., Ting, J.T., Wang, W., Wells, M.F., Venkatraman, T.N., Lascola, C.D., Fu, Z., and Feng, G. (2011). Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* *472*, 437-442.
128. Sperdin, H.F., and Schaer, M. (2016). Aberrant Development of Speech Processing in Young Children with Autism: New Insights from Neuroimaging Biomarkers. *Front Neurosci* *10*, 393.
129. Dinstein, I., Pierce, K., Eyler, L., Solso, S., Malach, R., Behrmann, M., and Courchesne, E. (2011).

- Disrupted neural synchronization in toddlers with autism. *Neuron* 70, 1218-1225.
130. Andersen, L.M., Jerbi, K., and Dalal, S.S. (2020). Can EEG and MEG detect signals from the human cerebellum? *NeuroImage* 215, 116817.
 131. Kikuchi, M., Shitamichi, K., Yoshimura, Y., Ueno, S., Remijn, G.B., Hirose, T., Munese, T., Tsubokawa, T., Haruta, Y., Oi, M., et al. (2011). Lateralized theta wave connectivity and language performance in 2- to 5-year-old children. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31, 14984-14988.
 132. Babiloni, C., Vecchio, F., Mirabella, G., Buttiglione, M., Sebastiano, F., Picardi, A., Di Gennaro, G., Quarato, P.P., Grammaldo, L.G., Buffo, P., et al. (2009). Hippocampal, Amygdala, and Neocortical Synchronization of Theta Rhythms is Related to an Immediate Recall During Rey Auditory Verbal Learning Test. *Human brain mapping* 30, 2077-2089.
 133. Hertrich, I., Mathiak, K., and Ackermann, H. (2016). The Role of the Cerebellum in Speech Perception and Language Comprehension. In *The Linguistic Cerebellum*. pp. 33-50.
 134. Basar, E. (2013). A review of gamma oscillations in healthy subjects and in cognitive impairment. *Int J Psychophysiol* 90, 99-117.
 135. Schwartz, M., and Kotz, S.A. (2016). Contributions of cerebellar event-based temporal processing and preparatory function to speech perception. *Brain and Language* 161, 28-32.
 136. Donovan, A.P., and Basson, M.A. (2017). The neuroanatomy of autism - a developmental perspective. *J Anat* 230, 4-15.
 137. Franchini, M., Zoller, D., Gentaz, E., Glaser, B., Wood de Wilde, H., Kojovic, N., Eliez, S., and Schaer, M. (2018). Early Adaptive Functioning Trajectories in Preschoolers With Autism Spectrum Disorders. *J Pediatr Psychol* 43, 800-813.
 138. Franchini, M., Wood de Wilde, H., Glaser, B., Gentaz, E., Eliez, S., and Schaer, M. (2016). Brief Report: A Preference for Biological Motion Predicts a Reduction in Symptom Severity 1 Year Later in Preschoolers with Autism Spectrum Disorders. *Front Psychiatry* 7, 143.
 139. Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Leventhal, B.L., DiLavore, P.C., Pickles, A., and Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders* 30, 205-223.
 140. Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., Bishop, S., Esler, A., Hus, V., Oti, R., et al. (2009). The Autism Diagnostic Observation Schedule-toddler module: a new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of autism and developmental disorders* 39, 1305-1320.
 141. Shank, L. (2011). Mullen Scales of Early Learning. In *Encyclopedia of Clinical Neuropsychology*, J.S. Kreutzer, J. DeLuca and B. Caplan, eds. (New York, NY: Springer New York), pp. 1669-1671.
 142. Boashash, B. (2015). *Time-frequency signal analysis and processing: a comprehensive reference*, (Academic Press).
 143. Pellegrino, F., Coupé, C., and Marsico, E. (2011). Across-language perspective on speech information rate. *Language* 87, 539-558.
 144. Jessen, S., and Kotz, S.A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *NeuroImage* 58, 665-674.
 145. Pichon, S., de Gelder, B., and Grezes, J. (2009). Two different faces of threat. Comparing the neural systems for recognizing fear and anger in dynamic body expressions. *NeuroImage* 47, 1873-1883.
 146. Sperdin, H.F., Coito, A., Kojovic, N., Rihs, T.A., Jan, R.K., Franchini, M., Plomp, G., Vulliamos, S., Eliez, S., Michel, C.M., et al. (2018). Early alterations of social brain networks in young children with autism. *Elife* 7.
 147. Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134, 9-21.
 148. Michel, C.M., and Brunet, D. (2019). EEG Source Imaging: A Practical Review of the Analysis Steps. *Front Neurol* 10, 325.
 149. Cohen, M.X. (2014). *Analyzing neural time series data: theory and practice*.
 150. Crosse, M.J., Di Liberto, G.M., Bednar, A., and Lalor, E.C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Front Hum Neurosci* 10, 604.
 151. Di Liberto, G.M., and Lalor, E.C. (2017). Indexing cortical entrainment to natural speech at the phonemic level: Methodological considerations for applied research. *Hearing Research* 348, 70-77.
 152. Di Liberto, G.M., Lalor, E.C., and Millman, R.E. (2018). Causal cortical dynamics of a predictive enhancement of speech intelligibility. *NeuroImage* 166, 247-258.
 153. Fiedler, L., Wostmann, M., Graversen, C., Brandmeyer, A., Lunner, T., and Obleser, J. (2017). Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed

- speech. *J Neural Eng* *14*, 036020.
154. Le Van Quyen, M., Foucher, J., Lachaux, J., Rodriguez, E., Lutz, A., Martinerie, J., and Varela, F.J. (2001). Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *J Neurosci Methods* *111*, 83-98.
 155. van Driel, J., Cox, R., and Cohen, M.X. (2015). Phase-clustering bias in phase-amplitude cross-frequency coupling and its removal. *J Neurosci Methods* *254*, 60-72.
 156. Combrisson, E., and Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *J Neurosci Methods* *250*, 126-136.
 157. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* *58*, 267-288.
 158. Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*.

δ  θ  β low- γ 

-3

-2

-1

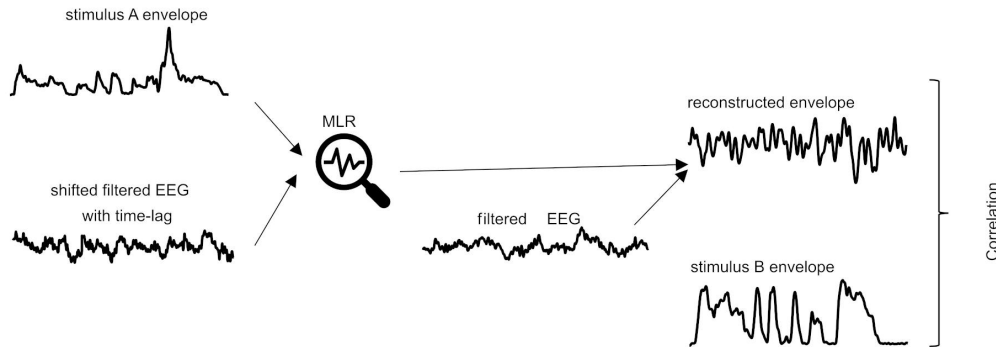
0

1

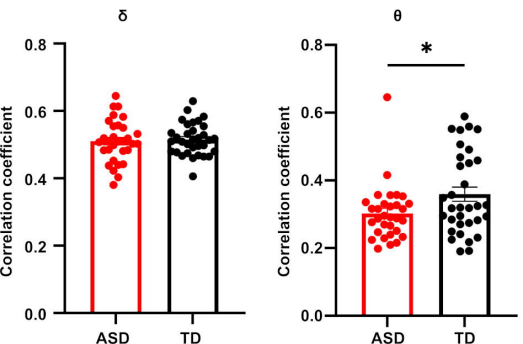
2

t-value

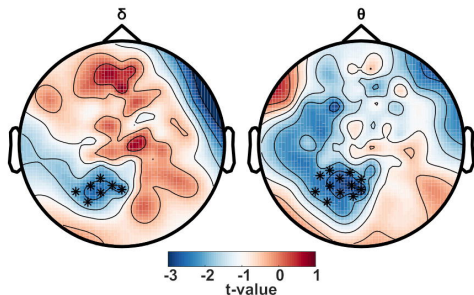
A. Schematics of the Multiple Linear Regression (MLR) model approach

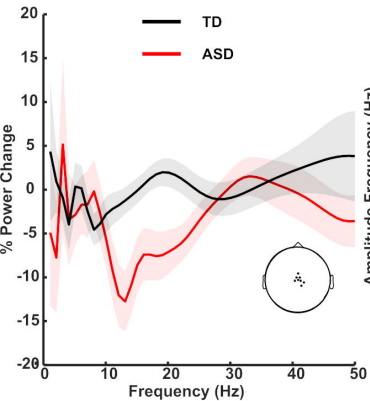
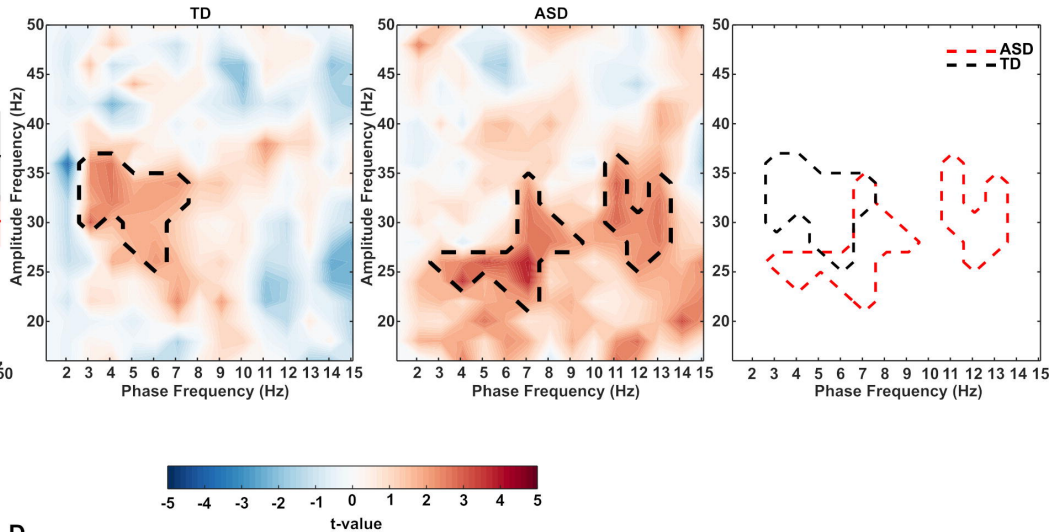
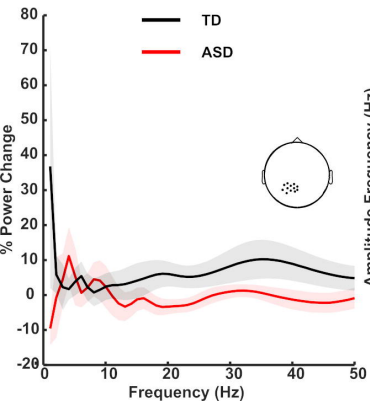
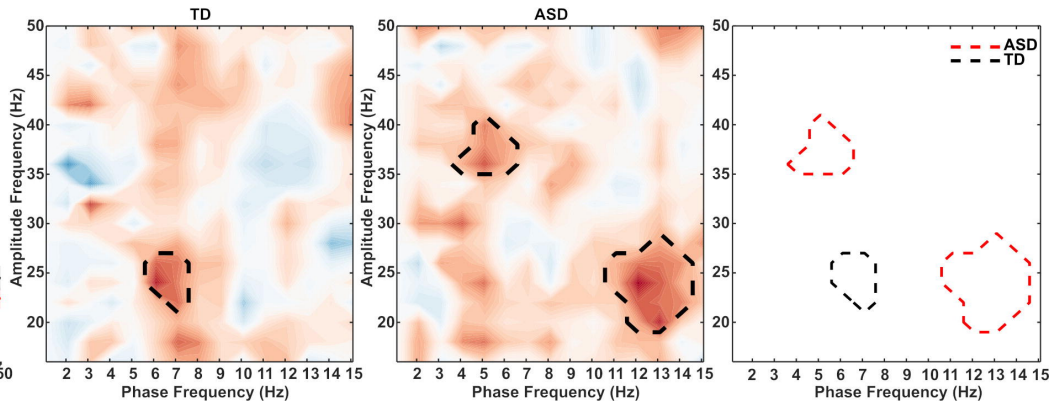


B. all electrodes analyses

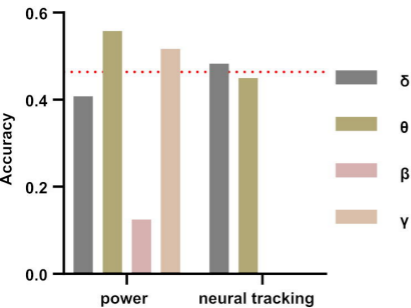


C. single electrode analyses

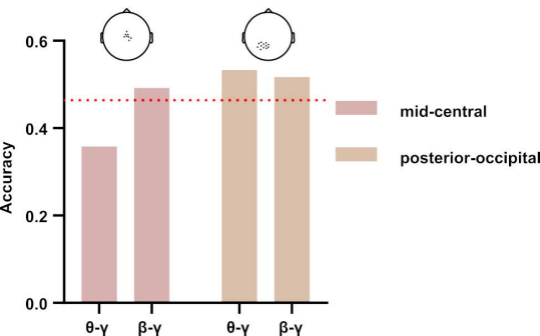


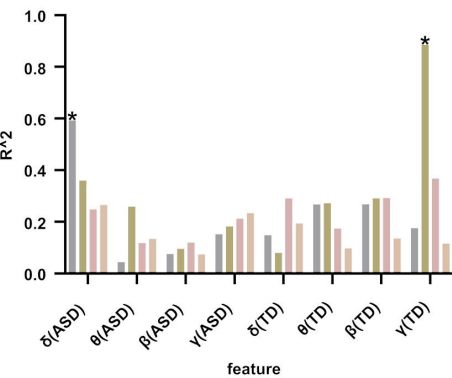
A.**B.****C.****D.**

A. EEG power and neural tracking

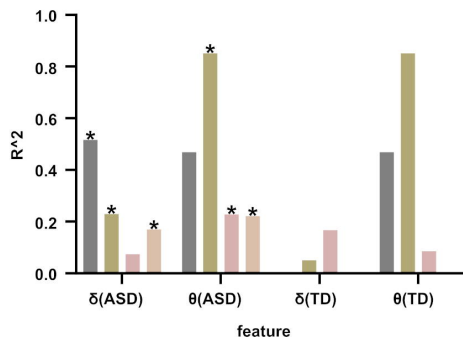
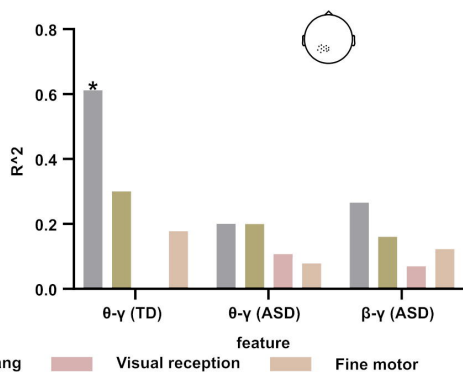
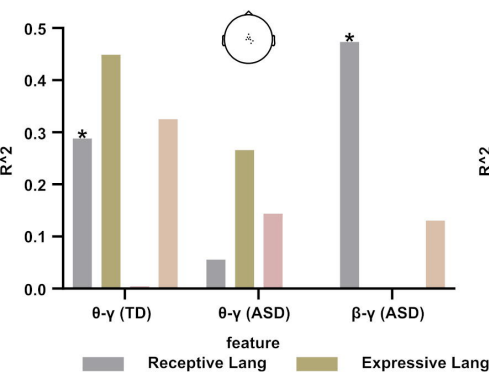


B. Phase-amplitude coupling



A. EEG Power

Receptive Lang Expressive Lang Visual reception Fine motor

B. Neural tracking**C. Phase-amplitude coupling (mid-central electrodes and posterior-occipital electrodes)****D. Language reception prediction from beta/gamma coupling**