

SCIENTIFIC REPORTS



OPEN

Language and music phrase boundary processing in Autism Spectrum Disorder: An ERP study

John DePriest^{1,2}, Anastasia Glushko^{1,4}, Karsten Steinhauer^{3,4} & Stefan Koelsch^{1,5}

Autism spectrum disorder (ASD) is frequently associated with communicative impairment, regardless of intelligence level or mental age. Impairment of prosodic processing in particular is a common feature of ASD. Despite extensive overlap in neural resources involved in prosody and music processing, music perception seems to be spared in this population. The present study is the first to investigate prosodic phrasing in ASD in both language and music, combining event-related brain potential (ERP) and behavioral methods. We tested phrase boundary processing in language and music in neuro-typical adults and high-functioning individuals with ASD. We targeted an ERP response associated with phrase boundary processing in both language and music – i.e., the Closure Positive Shift (CPS). While a language-CPS was observed in the neuro-typical group, for ASD participants a smaller response failed to reach statistical significance. In music, we found a boundary-onset music-CPS for both groups during pauses between musical phrases. Our results support the view of preserved processing of musical cues in ASD individuals, with a corresponding prosodic impairment. This suggests that, despite the existence of a domain-general processing mechanism (the CPS), key differences in the integration of features of language and music may lead to the prosodic impairment in ASD.

A large subgroup of individuals with autism spectrum disorder (ASD) shows deficient language abilities¹. One of the most prominent difficulties within the language domain is the impairment in prosodic processing^{2,3}, which has been directly linked to how social and communicative competence of individuals with ASD is assessed by others⁴.

The relevance of prosodic processing to the quality of communication is not only due to the affective functions of prosody⁵, but also to its close interaction with syntactic analysis^{6,7}. Several behavioral studies have looked at how prosodic features relevant for sentence parsing are processed in ASD^{4,8–14}. In behavioral tasks involving prosodic phrasing, or chunking, individuals with ASD have been shown to either not differ from their neuro-typical peers^{3,8,9,13,15} or, on the contrary, to indeed have difficulties in prosodic processing^{10,11}. The inconsistency of these findings may be partially attributed to a number of methodological challenges faced by behavioral studies on prosodic processing². More sensitive experimental approaches, such as neuroimaging, could, therefore, shed light on whether prosodic processing involved in sentence parsing is in fact impaired in ASD.

To date, few neuroimaging studies of prosodic processing in ASD have been conducted. An early functional Magnetic Resonance Imaging (fMRI) study found that affective prosody failed to recruit right-hemisphere regions in an ASD group, suggesting that this group relies on a different cortical network for prosodic processing¹⁶. Other fMRI studies, however, have examined the prosodic skills of high-functioning autistic individuals and found that while nearly identical right-hemisphere regions were activated in ASD and typically developing populations during both affective and linguistic prosody tasks, additional brain areas, including the left supra-marginal gyrus were recruited in the ASD group^{17,18}. The recruitment of these additional areas may reflect an overcompensation mechanism in which ASD individuals attend more closely to phonemic details¹⁷, perhaps due to less automaticity of language processing¹⁸. Another study, using near infrared spectroscopy, found that ASD subjects showed a similar right-hemisphere laterality of prosodic processing as neuro-typical subjects, despite demonstrating a prosodic impairment¹⁹. The authors theorized that the prosodic impairment may arise from the integration of acoustic features of language with semantic or syntactic factors. Together, these studies suggest that the neural substrates

¹Freie Universität Berlin, Berlin, Germany. ²Program in Linguistics, Tulane University, New Orleans, Louisiana, United States of America. ³School of Communication Sciences and Disorders, McGill University, Montreal, Quebec, Canada.

⁴The Centre for Research on Brain, Language and Music (CRBLM), Montreal, Quebec, Canada. ⁵University of Bergen, Bergen, Norway. Correspondence and requests for materials should be addressed to J.D. (email: jdepries@tulane.edu)

underlying affective and linguistic prosody in ASD are only partly overlapping with those in neuro-typical populations, and that the prosodic impairment may stem from integrating disparate features of language.

Whereas language impairment has been reliably observed in ASD, musical abilities in this group are consistently reported to be spared²⁰ or even heightened^{21–26}. Studies have also repeatedly provided evidence of overlapping, and possibly shared neurocognitive mechanisms underlying music and language perception in neurotypical populations²⁷. Language and music processing rely heavily on both left- and right-hemisphere structures, including superior temporal, and inferior fronto-lateral regions in both hemispheres, although syntactic information is usually processed with a left-hemisphere weighting in language, and right-hemisphere weighting in music^{28,29}. However, processing of prosodic features of language relies on right-hemisphere networks, similar to those responsible for tonal perception in music³⁰. Thus, language, and specifically prosodic impairment in ASD in the absence of a corresponding music impairment, presents a curious dissociation.

Established theories about the nature of ASD neurocognition fail to explain this dissociation. The underconnectivity framework has been proposed based on several studies demonstrating a decrease in structural long-range, primarily frontal-posterior, but also interhemisphere, connections between brain regions in the ASD population^{31–33}. The notion of long-range underconnectivity (and the resulting local over-connectivity) is in line with the behavioral findings of perceptual bias individuals with ASD have toward local, fine-grained characteristics of both auditory and visual input^{34,35}. Models such as Weak Central Coherence³⁵ (WCC) and Enhanced Perceptual Functioning (EPF)³⁶ have been proposed based on this observation. The former suggests that the local bias in ASD occurs as a consequence of an impairment in processing global characteristics of input. The EPF, in contrast, emphasizes that the focus on local information reflects superior low-level perceptual processing that is independent of the presence or the absence of difficulties in high-level processing. Additionally, the theory asserts that higher-level features of perception do not interfere in tasks that can be processed locally for an ASD population³⁶. However, WCC and EPF, as well as the corresponding proposal of long-range neural underconnectivity, fail to account for the observed discrepancy between language and music processing skills in ASD. Lai, Pantazatos, Schneider, and Hirsch²³ came to a similar conclusion: they showed that in children with autism, the dissociation between spoken language and song processing was characterized by differences in left inferior frontal gyrus (LIFG) activation compared to their neurotypical peers, rather than being marked by deviated connectivity patterns. They suggested instead that the preserved musical abilities may be attributed to differences between the low-level auditory cues essential for music, on the one hand, and language, on the other, (consistent with more recent findings^{37,38}), and that the language-specific low-level auditory information might not be received by the higher-level LIFG region due to an impairment in lower processing regions. Note, however, that Lai and colleagues²³ studied a group of low-functioning children with autism (compared to an age-matched group of neurotypical children), some of whom were sedated, which presents a challenge for drawing conclusions about the nature of the dissociation between language and music in autism.

To address the questions arising from the inconsistent behavioral reports of prosodic impairments and conflicting explanatory models of ASD, and to specify the relationship between prosodic and music processing in this disorder, we used event-related potentials (ERP). ERPs provide high-resolution information about the timing, and about the amount of neural resources of cognitive processing (making this technique sometimes more sensitive to subtle processing differences compared to behavioral tasks). With respect to prosody, ERPs offer a well-established paradigm to study prosodic chunking, which, importantly, can be analogously used for the study of musical phrasing. Neurophysiological studies of prosodic phrasing revealed the so-called Closure Positive Shift (CPS) – a positive-going electrophysiological response elicited by intonational phrase (IPh) boundary perception in both spoken^{39–43} and written language^{44,45}. IPhs are major prosodic chunks within a sentence, which normally correspond to syntactic phrases. IPh boundaries are marked by prosodic cues that may differ across languages and speakers, though in all cases, recognition of IPh boundaries plays an important role in the correct interpretation of a given sentence. The CPS is elicited immediately after the onset of the pause between two intonational phrases, can last for several hundred milliseconds, and typically has a distribution near the midline of the scalp⁴³. Whereas the CPS component is thought to index prosodic chunking of preboundary phrase elements in real time, mismatches between prosodic chunks and syntactic phrasing requirements in so-called ‘garden-path’ sentences may elicit a subsequent bi-phasic N400/P600 response on disambiguating words³⁹. This ERP pattern associated with processing of prosody-syntax mismatches provides an additional electrophysiological measure of successful prosody perception, independent of the CPS.

Several studies have also described a positive-going ERP waveform elicited close to phrase boundaries in music, which was said to resemble the CPS in language^{46–49}. This so-called ‘music-CPS’, however, is characterized by a latency and duration different from that found for the language-CPS. More recently, the functional significance of this post-boundary music-CPS has been drawn into question, and an additional CPS-like response, named the boundary-onset music-CPS, has been observed during the musical phrase boundary for both musicians and non-musicians⁵⁰. This response has a similar latency and distribution to the language-CPS, and is a prolonged positive shift. The similarity of these ERPs in language and music processing makes them useful tools for understanding differences between language and music cognition.

The present study is the first to investigate the CPS in response to both language and music in ASD. Importantly, the use of this experimental paradigm allowed us to make predictions disambiguating the different explanatory models proposed for the neurocognition in ASD in general and the dissociation between language and music skills in this group in particular. Based on existing research results, we have considered three possible outcomes. (1) The absence of differences in prosodic or musical phrasing between high-functioning adults with ASD and neurotypical individuals would be in line with the EPF model that does not imply an obligatory impairment in high-level cognitive processing mechanisms in ASD. (2) An impairment in both prosodic and musical phrasing reflected in the markedly different amplitude and/or latency of the two CPS components in ASD compared to those of the neurotypical group would be consistent with the general underconnectivity hypothesis

	Group	Verbal-IQ	Non-Verbal LPS	AQ	Handedness
Mean (SD)	NT	108.06 (14.5)	31.1 (4)	13.88 (4.75)	79.14 (46.2)
	ASD	114.36 (13.6)	30.4 (6.9)	35.8 (9.37)	56.1 (64.97)
<i>t</i>	NT v ASD	1.167	0.287	7.183	1.022
<i>p</i>	NT v ASD	0.255	0.778	<0.001***	0.322

Table 1. Results of the neuropsychological battery by group: Verbal-IQ from the MW-B (Lehrl 1995), Non-Verbal from section 4 of the Leistungsprüfungssystem (Horn 1962); a German version of the Autism Quotient (Baron-Cohen *et al.* 2001); a Handedness Questionnaire adapted from Oldfield (1971). Results stated in: Mean (with standard deviation in parentheses), except for *t* and *p* values from the *t*-test.

and WCC. (3) Impaired prosodic phrasing along with intact musical phrasing would imply that the difference between prosodic and musical phrasing skills in ASD stems not from a domain-general chunking impairment, but rather from language-specific difficulties, perhaps stemming from differences in low-level acoustic characteristics between language and music³⁷, such as the temporal complexity of the input²³, or the integration of multiple levels of information in language¹⁹. Both options (2) and (3) would imply that, as a consequence of the predicted prosodic impairment, we would also see a deficit in processing prosodically cued garden-path structures in the ASD group, resulting in no differences between ERP responses to sentences with or without a prosody-syntax mismatch.

Methods

Participants. Thirty-three participants were tested as part of the study. Five participants' datasets (two from the ASD group and three from the neurotypical group) were excluded from the analysis due to extensive EEG artifacts or due to being extreme outliers; the group descriptors below are representative only of the participants whose data were included in the analysis. Eleven participants (three women) had a diagnosis on the Autism Spectrum of High-Functioning Autism, Asperger's Syndrome, or both. Diagnoses were made by physicians according to the Autism Diagnostic Observation Schedule (ADOS)⁵¹ and Autism Diagnostic Interview - Revised (ADI-R)⁵², or based on the criteria of ICD-10⁵³ or the DSM-IV TR⁵⁴. These participants were recruited from the Berlin/Brandenburg metropolitan area either through psychiatrist recommendation, or through advertisements in organizations related to autism. The ages of the ASD participants ranged from 23 to 54 with a mean age of 37.45 (SD 9.5).

Seventeen neuro-typical (NT) participants (seven women) were included in the analysis. All participants gave informed consent and were paid for their participation. Ethics approval was obtained from both the FU Berlin and Tulane University, and all methods were performed in accordance with the relevant guidelines and regulations. The neuro-typical participants ranged in age from 21 to 56 with a mean age of 28.47 (SD 8.25). Although the difference in ages of the two groups was significant ($t = 2.571$, $p = 0.019$), the results were virtually identical when analyzing an age-matched subset of neuro-typical behavioral and EEG data. Further, no age effects have been observed in the previous language-CPS studies, and while the absence of N400 and a more anterior P600 have been correlated with normal aging, those results were from a much older group (65–80 years)⁵⁵. Thus, for purposes of statistical power, the larger group's data will be reported. Exclusion criteria for both groups comprised history of neurological problems (aside from ASD for the ASD group), a family member with ASD (for the neuro-typical group), or a history of alcohol or drug dependence.

Neuropsychological Battery. Participants were given a series of questionnaires, including the German translation of the Autism Questionnaire^{56,57}, the Multiple Choice Vocabulary Test-B (MWB) test for Verbal IQ⁵⁸, the Performance Testing System (LPS), Section 4 test for Non-Verbal IQ⁵⁹, a Handedness Questionnaire⁶⁰, and a brief in-house questionnaire assessing the musical background of participants.

The Autism-Spectrum Quotient (AQ) is a screening test designed to determine the degree to which an adult with normal intelligence has traits associated with the autistic spectrum, including assessments of social skills, attention to detail, and communication⁵⁶. A cutoff of score of > 32 was used to classify those individuals with a high degree of ASD traits. None of the neuro-typical participants scored above 23. Section 4 of the LPS is a logical, non-verbal reasoning test. Only raw scores from this section are shown. The results of the neuro-psychological battery can be found in Table 1.

Stimuli. The experimental paradigm was developed based on the one used by Steinhauer and colleagues³⁹ with language stimuli adapted from the study of Pannekamp and co-authors⁶¹ to make the task more appropriate for adult participants, and relied on the same stimuli used by Glushko and colleagues⁵⁰. The sentences were grammatically all of the same form (# indicates the IPh boundary):

Intransitive: 'Maxe bittet Tina zu lächeln, # und das Lied mitzusingen.'

Maxe asks Tina to smile, # and to sing a song.

Transitive: 'Maxe bittet, # Tina zu grüssen, # und das Lied mitzusingen.'

Maxe asks # to greet Tina, # and to sing a song.

These sentences differed primarily in the argument structure of the second verb. In the above examples, the second verb in the Intransitive condition, *lächeln* ('to smile'), is intransitive, while *grüssen* ('to greet'), in the

Figure 1. Sample notation of the music stimuli. Time periods of interest for the ERP analysis are highlighted with rectangles. Figure 1a: sample stimulus from the Phrased condition; Fig. 1b: Unphrased condition; Fig. 1c: No Pause condition.

Transitive condition is transitive and takes ‘Tina’ as its accusative argument. The sentences follow the same word order, yet the intonation patterns with which these two sentences are spoken differ, aiding in the hearer’s syntactic parsing. Pauses at the IPhs were silenced, and edited for consistency to a duration of 600 ms.

In addition, we constructed prosodically and syntactically incongruous sentences (the ‘Mismatch’ condition) to test the validity of any prosodic phrase boundary processing impairment. If there are differences between the two groups’ language-CPS, we wanted to ensure that they were due to the impairment in processing phrase boundaries in language (shown by diminished garden-path effects), rather than more efficient processing of intonational phrase boundaries (shown by comparable garden-path effects). Thus, the Transitive and Intransitive sentences were carefully cross-spliced during the alveolar closure at the beginning of the affricate /ts/ in the word ‘zu’, so that the phrase up to that point came from the Transitive condition, while the phrase following that point was from the respective Intransitive sentence. This resulted in prosodically and syntactically incongruous sentences, such as:

Mismatch: *‘Maxe bittet, # Tina zu lächeln, # und das Lied mitzusingen.
*Maxe asks, # to smile Tina, # and to sing a song.

In total, 48 sentences of each of the three conditions were produced, resulting in a total of 144 sentences.

The music stimuli were composed for this project by the first author, who is also a professional musician, and reviewed by a professional composer. The samples were composed in Sibelius First (Version 6.1.5, Avid Technology, Burlington, USA) and exported with a realistic acoustic piano sound. Each track included 16 bars at 100 beats per minute (bpm) with an upbeat. Each musical sample followed the same structure of two four-bar phrases creating an 8 bar piece, which was repeated. Three musical conditions were created for this study (Fig. 1):

Phrased (Fig. 1a): a pre-boundary half note followed by a pause (a quarter note rest lasting 600 ms);

Unphrased (Fig. 1b): a pre-boundary half note followed by either a quarter note or two eighth notes (also 600 ms);

No Pause (Fig. 1c): a boundary dotted half note of 1800 ms duration with gradual amplitude decay.

The names of the first two categories are used by convention, even though as Nan and colleagues⁴⁸ acknowledge, the Unphrased condition is in fact ‘less phrased’, since other phrasing cues remain in addition to a pause or rest, including a lengthened phrase-final note, harmonic closure, and the structural convention of 4-bars. The third condition was designed to determine if a musical CPS associated with musical phrase boundaries is dependent on the presence of a pause (or rest). Further, previous studies had not examined the time window between phrases, which corresponds more closely to where the language-CPS occurs. However, in the absence of this third condition, the comparison of responses in this time window would be confounded by the presence of silence in one condition and additional notes in the other. No ‘Mismatch’ condition was used in the music part of

Effect	Language		Music	
	CPS	N400/P600	Boundary-onset music-CPS	Post-boundary music-CPS
Baseline (ms)	−500 to 0; −50 to +50	None: peak-to-peak	−2000 to −1800; −1800 to −600	−2000 to −1800*
TWs (ms)	0 to +600	Min: +200 to +650 max: +600 to +1200	−450 to 0	+450 to +600

Table 2. Summary of the time windows (TW) and baseline corrections used for analysis of EEG data. All values expressed in milliseconds (ms). *A baseline-independent analysis was also used for this TW. See Supplementary Materials for details.

the experiment due to the lack of a musical equivalent to the prosody/syntax mismatch. Importantly, the elicitation of the (language-)CPS is virtually independent of the presence (or absence) of violations or anomalies in the sentence⁴⁴ and task requirements³⁹. Therefore, whether a Mismatch condition was present in speech and not in music should have little to no impact on the elicitation of the CPS components.

Procedure. All measurements were performed at the Dahlem Institute for the Neuroimaging of Emotion Laboratory at the Freie Universität Berlin. The stimuli were presented using Presentation[®] software (Neurobehavioral Systems, Berkeley, USA) in separate counterbalanced language and music blocks, with stimuli presented in randomized order. The entire procedure, including the completion of consent forms and questionnaires, lasted approximately 120 minutes. Following the presentation of each sentence, participants were asked “How natural did you find the last sentence?” Responses were provided on a five-point scale from “completely unnatural” to “completely natural”. This task mainly sought to determine if the ASD group was able to accurately perceive the violation in the Mismatch condition. The corresponding task for the music stimuli was to answer the question “How natural did you find the last piece of music?”, the main purpose of which was to maintain participant attention during stimulus presentation.

EEG Recording and Analysis. All EEG recordings were conducted using 32 Ag/AgCl electrodes distributed according to the international 10–20 system using Brain Vision Recording Software (Brain Products GmbH, Gilching, Germany) at a sampling rate of 500 Hz. Electrodes were attached on both mastoids, of which the right mastoid was used as a reference electrode. Also, a ground was attached to the nape of the neck, and all impedances were kept below 10 kΩ.

EEG recordings were analyzed using EEProbe software (ANT Software BV, Enschede, Netherlands). Eye-blinks and facial movements were then corrected using a regression-based statistical algorithm, or rejected manually. Participants with more than 40% of trials rejected in any of the conditions (one participant with ASD and two neurotypical participants) were not included in the analysis. A band-pass FIR filter of 0.3–30 Hz (1001 points) was used to remove line noise and artifactual slow waves.

In the language stimuli, epoch measurements were time-locked to the beginnings of pauses marking phrase-boundaries. For CPS analyses in both language and music, multiple baseline correction regions were used to compare the stability of effects observed during different conditions (see Table 2). To quantify the CPS component, ERP responses were then measured across the 600 ms time windows (TWs) following the phrase boundaries. In the Mismatch condition, the prosody-syntax mismatch effects were assessed time-locked to the onset of the disambiguating second verb of the sentences. Intransitive and Mismatch conditions were contrasted using peak-to-peak analysis preceded by filtering of the EEG data with a 5 Hz low-pass filter. Local minima and maxima were identified for the time windows of 250 to 650 ms and 600 to 1200 ms, respectively, at electrodes Cz and Pz following the onset of the second verb. These electrodes were chosen due to the central and parietal distribution of the N400/P600 responses to garden-path sentences in our and previous studies^{39,50}.

All epochs and baselines in the music experiment are stated in reference to the onset of the first post-boundary note (see Table 2). TWs for analysis included: the 450 ms preceding the onset of the post-boundary phrase-initial note for the boundary-onset music-CPS, and between 450 and 600 ms following the onset of this note (for the post-boundary music-CPS: see Supplementary Materials).

Statistical Analysis. Statistical analysis of behavioral data was conducted using the statistical analysis software *R* (*R* Foundation for Statistical Computing, Vienna, Austria), and SPSS (Version 22, IBM, Armonk, USA). Behavioral responses were averaged by participant and condition, and then analyzed using two repeated measures 2×3 analyses of variance (ANOVA) with the two level between-subject factor Group (ASD vs. NT), and the three level within-subject factor Condition composed of the three language (Intransitive vs. Transitive vs. Mismatch) and music (Phrased vs. Unphrased vs. No Pause) conditions, respectively. Post-hoc comparisons between condition pairs for each group were performed with Two-Sample *t*-tests, using Bonferroni correction for multiple comparisons.

EEG data were grouped into regions of interest (ROIs) and averaged across electrodes for each condition and participant. Lateral ROIs included 18 electrodes, with 3 in each ROI: AntLeft: F3, F7, FC5; AntRight: F4, F8, FC6; CentLeft: C3, CP5, T7; CentRight: C4, CP6, T8; PostLeft: P3, P7, O1; and PostRight: P4, P8, O2. Midline comparisons were conducted using electrodes Fz, Cz, and Pz. Any participant whose average ERP response (lateral or midline) in any condition, language or music, fell more than ± 2.5 Standard Deviations away from the mean of the group was considered as an outlier and excluded from analysis. Then, repeated measures ANOVAs were conducted for all TWs for both lateral and midline ROIs. For the lateral electrodes, ANOVAs containing

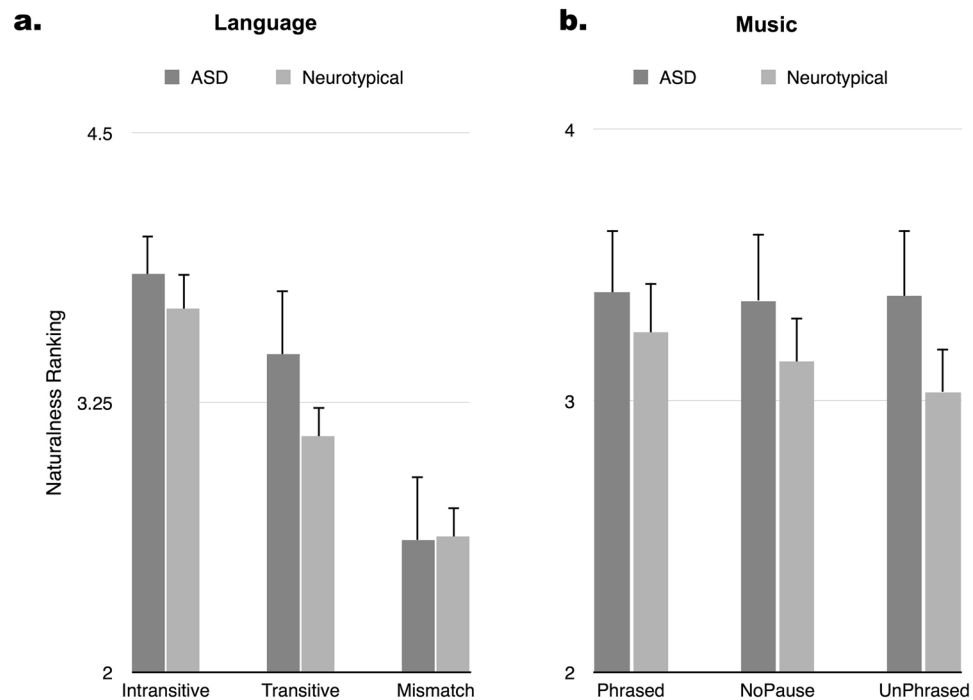


Figure 2. Results of the behavioral experiment. Figure 2a: mean naturalness ratings (1 – least natural, 5 – most natural) with standard error bars of the sentences in the language part of the experiment by group. There were no between-group differences in naturalness ratings, but both groups rated the Mismatch condition as the least natural and participants also rated the Transitive condition as less natural than the Intransitive condition. Figure 2b: responses to the question “how natural did you find the last piece of music?” (1 – least natural, 5 – most natural). There were no differences observed by condition or group.

the factors Condition (language-CPS TW: Intransitive vs. Transitive; N400/P600 TW: Intransitive vs. Mismatch; Boundary-onset music-CPS TW: Phrased vs. No Pause), Hemisphere (L vs. R), AntPost (Ant vs. Cent vs. Post), as well as the between-subject factor of Group (ASD vs. NT) were conducted. For the Midline electrodes, similar repeated measures ANOVAs were conducted without the factor Hemisphere and using electrode site for the factor AntPost (Fz vs. Cz vs. Pz). Greenhouse-Geisser corrections for sphericity were applied where appropriate. Uncorrected degrees of freedom are shown accompanied by corrected p values. Effect sizes are reported for all main effects and interactions using Cohen’s d .

Due to the widely recognized heterogeneity of ASD, and the relatively small sample size, each EEG data set was averaged across lateral and midline electrodes, respectively, and then tested for normal distribution using Kolmogorov-Smirnov distribution tests. This resulted in normal distributions for all analyses with the exception of the Midline analysis for the boundary-onset music-CPS data subset. Therefore, for this analysis, non-parametric Wilcoxon Signed Rank and Mann-Whitney U tests were used instead of ANOVA.

Results

Behavioral Results. For the language responses (Fig. 2a), the Intransitive condition was ranked as most natural (Mean: 3.749, SD: 0.66), while the Transitive sentences (Mean: 3.244, SD: 0.67) were ranked as less natural than the Intransitive ones. The Mismatch condition, which contained the violation, was ranked least natural of the three (Mean: 2.622, SD: 0.79). Main effects of Condition were significant at $p < 0.001$ for the global ANOVA ($F(2, 25) = 32.874, p < 0.001$) and for all t -tests following Bonferroni correction (Intransitive vs. Transitive: $t(27) = 6.121, d = 0.763$; Intransitive vs. Mismatch: $t(27) = 6.536, d = 1.55$; Transitive vs. Mismatch: $t(27) = 4.283, d = 0.852$). No significant main effects of Group were observed, nor was an interaction of Condition \times Group. The lack of between-group differences was not entirely unexpected, given that previous studies have reported inconsistent results of behavioral experiments on prosody (one of the motivations for the use of ERPs).

For the ratings of the music stimuli (Fig. 2b), no main effects nor interactions of Group nor Condition were observed, however, none of the musical stimuli contained violations, and only differed in the degree of phrasing. Because the purpose of the behavioral task in the music conditions was to maintain the attention of the participants, these results are consistent with our expectations. Overall, no between-group differences were observed when behavioral measures were used.

EEG Results: Language. The EEG results of the language task are summarized in Figs 3 and 4. For the sake of brevity, only the results of the -500 to 0 baseline correction for the first CPS are reported, except when meaningful differences were seen between baseline corrections.

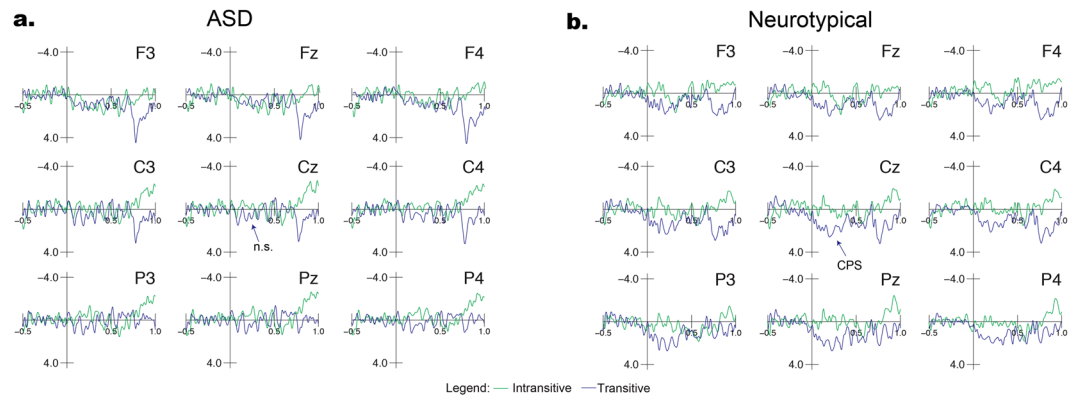


Figure 3. Brain-electric responses (group averages) to the Intransitive (without IPh boundary) and Transitive (with IPh boundary) conditions using a baseline of -500 to 0 ms. Figure 3a shows ASD group responses, while Fig. 3b shows neurotypical group responses. A clear CPS was observed in the Neurotypical group in response to the Transitive condition, while differences between conditions were not significant for the ASD group.

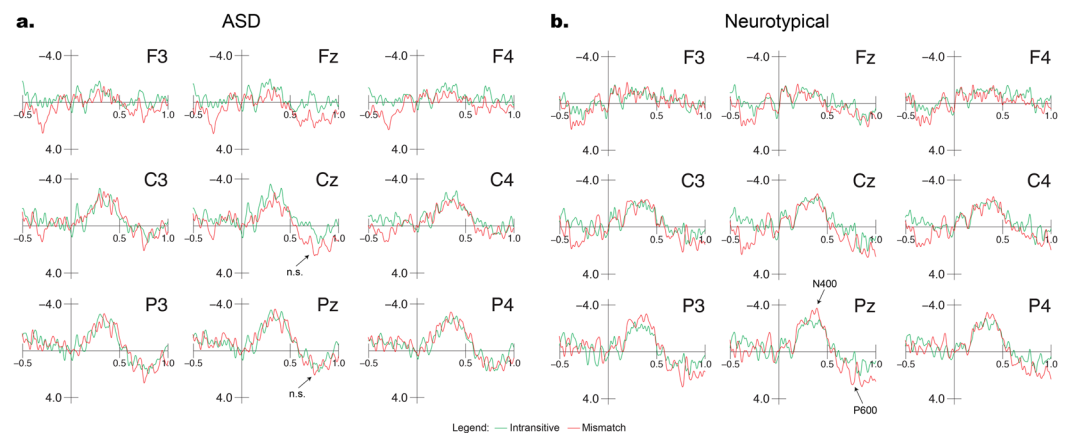


Figure 4. Brain-electric responses (group averages) to the Intransitive and Mismatch comparisons during the TW of expected Garden Path effects. Baseline correction of -50 to $+50$ ms was used for demonstration purposes, while peak-to-peak comparisons were used for analysis. Figure 4a shows ASD group responses which did not differ by Condition. Figure 4b shows neurotypical group responses which contain Garden Path (N400/P600) effects.

Language-CPS. As can be seen from the EEG waveforms in Fig. 3b, a clear language-CPS was observed in the NT group in response to the Transitive condition. This positivity started at the onset of the boundary pause and had a broad centro-parietal scalp distribution. The clear positivity for the NT group contrasted with the less distinct positivity with a smaller amplitude and more limited distribution for the ASD group (Fig. 3a).

A significant main effect of condition was observed for ANOVAs of both lateral ($F(1, 26) = 11.619$, $p = 0.002$, $d = 0.883$) and midline ($F(1, 26) = 7.807$, $p = 0.01$, $d = 0.771$) electrodes between the Intransitive and Transitive conditions, confirming the interpretation based on visual inspection of the language-CPS. The Condition \times Hemisphere interaction ($F(1, 26) = 3.648$, $p = 0.067$, $d = 0.197$) approached significance, suggesting that the language-CPS responses were slightly larger over the right-hemisphere. Moreover, a Condition \times Group interaction was found for comparisons at both lateral ($F(1, 26) = 5.163$, $p = 0.032$, $d = 0.515$) and midline ($F(1, 26) = 5.192$, $p = 0.031$, $d = 0.535$) sites. Post-hoc analysis conducted for each group showed that for the NT group, there was a highly significant effect of Condition with a large effect size (lateral: $t(16) = 5.138$, $p < 0.001$, $d = 1.388$; midline: $t(16) = 4.567$, $p < 0.001$, $d = 1.34$), while for the ASD group, the effect of Condition was not significant and the effect size was small (lateral: $t(10) = 0.626$, $p = 0.545$, $d = 0.259$; midline: $t(10) = 0.285$, $p = 0.781$, $d = 0.118$). That is, despite the apparent presence of a slight positive response to the Intransitive condition for the ASD group, this effect was not significant. Additionally, we performed simple linear regressions to investigate whether age and VIQ scores of participants in each of the experimental groups could predict the variability of the language-CPS amplitude at midline electrodes, but none of the results was statistically significant.

Garden-Path Effects. Figure 4 shows the grand average waveforms for the ASD (Fig. 4a) and NT (Fig. 4b) groups for the Intransitive and the Mismatch conditions, which are acoustically identical during this period.

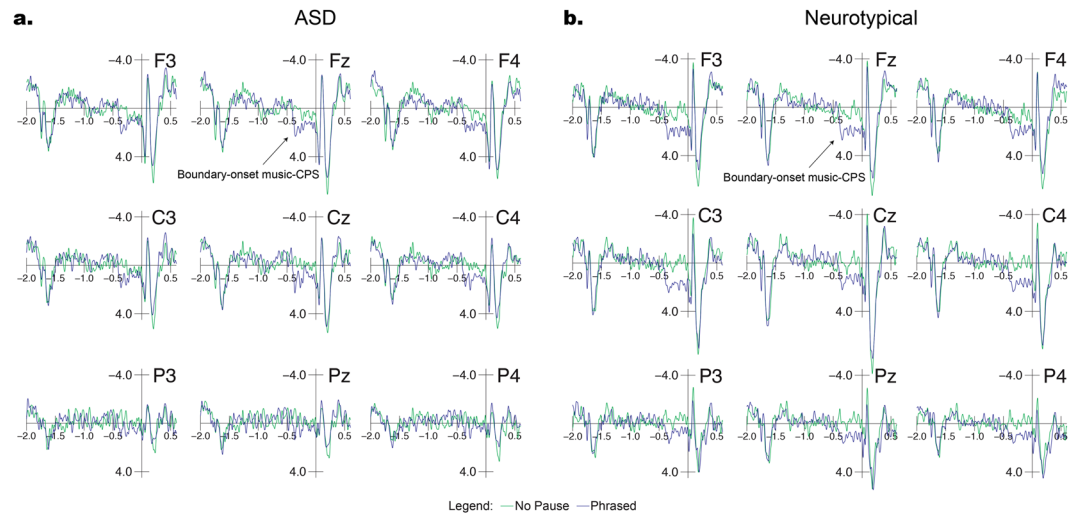


Figure 5. Brain-electric responses (group averages) to the boundary-onset music-CPS using a baseline of -1800 to -600 ms relative to the onset of the post-boundary phrase. Figure 5a shows ASD group responses, while Fig. 5b shows Neurotypical group responses. Both groups show a clear boundary-onset music-CPS in response to the Phrased condition.

The baseline of -50 to $+50$ ms relative to the onset of the verb is used for demonstration purposes (uncorrected peak-to-peak comparisons were used for statistical analysis). While no effects were significant at electrode Cz, a main effect of Condition (Intransitive vs. Mismatch) approached significance at electrode Pz ($F(1, 26) = 4.193$, $p = 0.051$, $d = 0.466$), as did the interaction of Condition \times Group at the same electrode ($F(1, 26) = 3.591$, $p = 0.069$, $d = 0.301$). Because a detailed comparison of the violation responses was crucial to examining ASD prosodic processing, follow-up pairwise t -tests were conducted to determine if garden-path responses were present for both groups. These pairwise t -tests conducted for the two conditions showed that for the NT group, there was a significant effect of Condition with a large effect size ($t(16) = 2.727$, $p = 0.015$, $d = 0.81$) due to an N400/P600 response to the Mismatch condition, while for the ASD group, the difference in responses to the two conditions was not significant and the effect size was negligible ($t(10) = 0.143$, $p = 0.889$, $d = 0.026$). We also performed simple linear regressions to investigate correlations between garden-path effects at Pz for both groups with age or VIQ score, but none of the results reached statistical significance.

EEG Results: Music. All TWs and baselines are stated in milliseconds relative to the onset of the first note following the phrase boundary in the Phrased and No Pause conditions, and the corresponding note in the Unphrased condition.

Boundary-Onset Music-CPS. The results of the two baseline corrections were nearly identical; therefore, we report the results using a -1800 to -600 ms baseline, mentioning the results of the -2000 to -1800 baseline only in cases of meaningful differences between the two analyses. Due to the presence of confounding notes ‘filling in’ the phrase boundary in the Unphrased condition, the analysis of this TW focused on comparisons of the Phrased and No Pause conditions.

As can be seen from the comparison of these two conditions in Fig. 5, both groups showed a positive shift during the phrase boundary in response to the Phrased condition. This shift resembled the language-CPS, as it was a prolonged positivity occurring during the pause between two phrases, with a central maximum. Overall, however, this response had a slightly more anterior distribution than the language-CPS in the present study (at the same time, note that a fronto-centrally distributed language-CPS has previously been reported by several studies^{42,43,61,62}).

In the TW of -450 to 0 , the two condition (Phrased vs. No Pause) ANOVA showed significant main effects of Condition for both lateral ($F(1, 26) = 15.739$, $p = 0.001$, $d = 0.821$), and midline comparisons (Wilcoxon Signed Rank $p = 0.001$, $d = 0.933$). This effect appeared smaller in the ASD group through visual inspection of the waveforms, yet, importantly, there was no significant Condition \times Group interaction at lateral sites ($F(1, 26) = 1.615$, $p = 0.215$, $d = 0.246$) and there were no significant differences between groups for either condition at midline sites (Mann Whitney U Phrased: $p = 0.306$, $d = 0.416$; Mann Whitney U No Pause: $p = 0.578$, $d = 0.065$). There were also significant main effects of Condition for the NT group (Wilcoxon Signed Rank NT $p = 0.001$) and the ASD group approached significance (Wilcoxon Signed Rank ASD $p = 0.05$). Additionally, there was a significant Condition \times AntPost interaction at lateral sites ($F(1, 26) = 5.864$, $p = 0.01$, $d = 0.75$), confirming the more anterior distribution of effects.

Discussion

Behavioral Results. We found no behavioral differences between groups in either the music or language tasks. Both the NT and ASD groups rated the (language) Mismatch condition as the least natural. The absence of group differences was somewhat contrary to our expectations: reconciling prosodic and syntactic information

from disparate processing pathways in ASD is counter to predictions based on underconnectivity, and Weak Central Coherence (WCC). However, both groups also rated the Transitive condition as less natural than the Intransitive condition. A possible reason for this could have been the silence during the pause at the first IPh of these sentences. The consistent, relatively long (600 ms) duration of the pause at the first IPh in the Transitive and Mismatch conditions may have influenced naturalness rankings by making the sentences appear edited. If so, the violation present in the Mismatch condition may not have been the only factor causing the groups to rank it as least natural. However, many sentences prior to editing already contained pauses of approximately the same duration, and other language-CPS studies⁶¹ have used stimuli with comparable IPh pause lengths. As mentioned earlier, behavioral measures of prosody in ASD have resulted in inconsistent findings². Further, behavioral rankings may reflect a variety of cognitive mechanisms, which was one of the motivating factors in our decision to use the more sensitive ERP technique to study prosody in an ASD group.

Language-CPS and Garden-Path Effects. A language-CPS was clearly present for the NT group in response to the Transitive condition, which contained an IPh boundary (as compared with the Intransitive condition, which did not). In line with previous studies^{39,43}, the language-CPS was elicited almost immediately following the onset of the pause, and was characterized by a broad centro-parietally maximal distribution near midline electrodes, lasting between 300 and 500 ms, which was somewhat larger over the right hemisphere. These findings are consistent with those of previous language-CPS results in distribution, duration, and amplitude^{39,43}. There was also a significant interaction of Condition \times Group and follow up tests showed no differences in responses to the different conditions for the ASD group. While the ASD group appeared to show a slight CPS in response to the Transitive condition, this response was not significant. The lack of a language-CPS response for this group shows that prosodic chunking ability is impaired in individuals with ASD, consistent with the perceptual prosodic challenges often noted by individuals with these disorders^{2,3}. While previous behavioral results have been inconsistent in describing the prosodic phrasing impairment in ASD, the use of more sensitive ERP measures in the present study confirms the presence of a linguistic prosody impairment for this group. The present study is the first to measure CPS data in ASD individuals, and to provide electrophysiological evidence of a prosody-based chunking impairment in this group.

Significant bi-phasic N400/P600 responses for the neuro-typical group were identified following the incongruous syntactic/prosodic splicing in the Mismatch condition, again replicating previous findings of similar experiments³⁹. By contrast, no garden-path ERP effects were observed in the ASD group. The absence of violation responses in this group suggests that the differences between the Mismatch rankings and those of the other conditions in the behavioral results may have been partially due to factors other than recognition of the prosody-syntax mismatch, such as the presence of an artificially silenced pause with a consistent duration. The inconsistency of the behavioral measures with the ERP data confirms one of our motivations for the study: that more sensitive electrophysiological measures can provide new insights in the investigation of prosody in ASD. Further, the absence of garden-path responses in this group is in line with previous research reporting N400 and P600 differences in ASD. Braeutigam and colleagues⁶³, summarizing previous research, found that ASD N400 responses vary substantially by task, with some studies finding no N400 effect, and others finding delayed effects. Recent studies have also failed to find a P600 response unless ASD participants were expressly told to focus on the degree of semantic implausibility in the sentences presented^{64,65}. These previous studies demonstrate broader patterns of language pathology in ASD; in the present study, the absence of N400/P600 responses can have a twofold interpretation. First of all, given the non-significant language-CPS response to transitive compared to intransitive sentences in participants with ASD, the absence of a typical N400/P600 pattern in sentences with a syntax-prosody mismatch is not surprising and primarily reflects participants' failure to process prosodic boundaries (i.e., leading to their inability to detect the associated syntax-prosody mismatch). In addition to this, it would be plausible, based on the underconnectivity hypothesis, to assume that the prosodic chunking impairment is accompanied by a distinct inability to integrate prosodic and syntactic information within a sentence. However, whether such an integration problem is present in ASD remains an open question.

Boundary-Onset Music-CPS. The present study is the first to identify a positive shift lasting several hundred milliseconds during musical phrase-boundaries for individuals with ASD. This response did not statistically differ between groups. While the distribution of this boundary-onset music-CPS was more anterior than the language-CPS, its time-course was more similar to the language-CPS than that of the previously described post-boundary music-CPS, which we and several previous studies failed to find^{66,67} (for reasons why the interpretation of a post-boundary music-CPS may be problematic, see Supplementary Materials and Glushko *et al.*⁵⁰; see also the absence of such an effect in the group of non-musicians in the study by Neuhaus and colleagues⁴³). The boundary-onset music-CPS reported here and elsewhere^{50,68} occurs during the phrase boundary, as opposed to 500-600 ms following the first post-boundary note. Additionally, the fact that it is a relatively slow positive shift, as opposed to being a single peak-like positivity, suggests that this component is the musical equivalent to the language-CPS, which in some previous studies showed a comparable frontal distribution^{42,61,69}. Together, these two responses reflect a domain-general neurocognitive mechanism of chunking⁵⁰.

In both language and music, the acoustic cues signaling phrase boundaries have similar characteristics. The phrase final note or syllable tends to be lengthened, accompanied by a change in pitch, and followed by a pause. In music, there are the additional features of strong beats, as well as other components of metrical hierarchies⁷⁰, that influence probable locations of phrase boundary divisions. The observed boundary-onset music-CPS likely represents the chunking of melodic or rhythmic phrases based on these cues, similar to the language-CPS. The fact that the positive shift was only observed in the condition with a musical pause, as opposed to the condition in which the phrase final note was held, suggests that the pause is crucial for the chunking of musical phrases.

Whether the boundary-onset music-CPS can be elicited in the absence of a musical pause (and, if so, under which conditions) remains unresolved. Note, however, that in language, where the CPS closely follows the morphology of the boundary-onset music-CPS, the response is produced at phrase boundaries even in the absence of a pause, provided that other phrase final characteristics are present³⁹.

Dissociation between music and language processing in ASD. In the present study, we reported impaired processing of prosodic phrase boundaries along with preserved processing of musical phrase boundaries in ASD. Despite the CPS components, observed in both language (language-CPS) and in music (boundary-onset music-CPS), reflecting a domain-general neurocognitive mechanism of chunking⁵⁰, the discrepancy between the elicitation of these ERPs in ASD presents a compelling dissociation.

The dissociation in the elicitation of the CPS in these two domains could plausibly stem from the inability of the higher-level chunking mechanism to access and rely on the lower-level perception of phrase boundary cues. In music, fine-grained pitch discrimination is crucial for the perception of harmony and consonance³⁷. The ability of individuals with ASD to process fine-grained, local information, such as pitch and note changes within a melody (for a review, see³⁵), would also allow for effortless pause perception at musical phrase boundaries. Therefore, music-specific phrase boundary processing (as reflected by the boundary-onset music-CPS) would be unimpaired.

In contrast to music, prosodic input presents a known processing difficulty for individuals with ASD¹ and calls for increased activation of areas responsible for perception of low-level features of language (compared to the neurotypical population)¹⁸. While music processing as a whole relies on fairly exact perception of pitch changes, for language sentence prosody processing, coarser pitch change perception is sufficient³⁷ and, in fact, might be essential for adequate speech analysis. The over-focus on fine-grained input features observed in individuals with ASD^{34,35} might actually be disadvantageous for their language processing abilities, driving attention away from prosody towards more rapid changes in phonemic information¹⁷ (which are not typical for music, see also³⁸). Language perception relies on coordinating distinct slow (i.e. prosodic) and more fine-grained (i.e. phonemic) levels for successful parsing³⁸, suggesting that an impairment in linking the two levels could also be responsible for the observed prosodic impairment. This interpretation of the dissociation between prosodic and musical phrasing in ASD is in line with the EPF model, which sees the local bias as a primary (rather than as secondary to the weak coherence) characteristic of ASD³⁶. It also potentially fits the idea of local neural over-connectivity in ASD^{31–33} that would imply a right-hemispheric local over-connectivity allowing for proper harmony²⁸ and fine-grained pitch^{71,72} processing as well as left-hemispheric local over-connectivity producing the bias towards phonemic analysis¹⁷. While it seems that the presence of the boundary-onset music-CPS in ASD would not easily fit predictions of underconnectivity, if the impairment caused by underconnectivity is one of integrating low-level and higher-level information together, this could explain the observed dissociation between language and music abilities in ASD. The possibility that superior pitch discrimination skills serve as an overcompensation mechanism boosting initially deficient phrasing abilities of individuals with ASD in the music domain, while the same overcompensation would not be successful in the language domain due to the relative independence of phonemic and prosodic levels, remains open and requires further investigation. It is clear, however, that in the language domain, individuals with ASD experience difficulties in prosodic chunking and these difficulties negatively affect their language comprehension skills (as reflected in the absent response to syntax-prosody mismatches in our study). This result is in line with previous behavioral reports of impaired prosodic chunking and descriptions of general linguistic abilities in ASD^{2,10}.

Conclusions

The present study is the first to investigate the electrophysiological responses to prosodic and music phrase-boundaries in individuals with autism spectrum disorders (ASD). While previous behavioral studies of ASD prosodic abilities have been contradictory and inconclusive, the use of the ERP technique strongly suggests a prosodic processing impairment in this group, with no corresponding impairment in music. In fact, the presence of nearly identical responses to music from the two groups provides empirical evidence that musical phrase structuring ability is preserved in ASD. This is also the first study to show a positive shift during the musical phrase boundary in individuals with ASD, similar to that found in neuro-typical individuals. This positivity may serve a similar function to the language-CPS, indexing a chunking mechanism that aids hierarchical and metrical structuring of musical input into cohesive phrases, while making predictions about what will come next. This interpretation implies that the dissociation between music and language processing abilities, particularly phrase boundary recognition, in ASD, stems from the differences between the acoustic properties of language and music and is potentially due to the allocation of neurocognitive resources to features of language other than prosody.

Both language and music abilities rely on the integration of overlapping networks incorporating various left and right hemisphere structures responsible for acoustic perception and rule-based operations, yet in ASD language ability is impaired, while music ability is not. Further research is needed to address the interaction of low-level and hierarchical features of music, as well as therapeutic avenues for reducing prosodic impairment, including whether musical training plays a role in improving prosodic processing and production, and if strategies for identifying prosodic phrase boundaries can decrease prosodic ambiguities in ASD.

References

1. Kjelgaard, M. M. & Tager-Flusberg, H. An investigation of language impairment in autism: Implications for genetic subgroups. *Lang. Cogn. Process.* **16**, 287–308 (2001).
2. McCann, J. & Peppé, S. Prosody in autism spectrum disorders: a critical review. *Int. J. Lang. Commun. Disord.* **38**, 325–350 (2003).
3. Peppé, S., McCann, J., Gibbon, F., O'Hare, A. & Rutherford, M. Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *J. Speech Lang. Hear. Res.* **50**, 1015 (2007).

4. Paul, R. *et al.* Brief Report: Relations between Prosodic Performance and Communication and Socialization Ratings in High Functioning Speakers with Autism Spectrum Disorders. *J. Autism Dev. Disord.* **35**, 861 (2005).
5. Nygaard, L. C. & Queen, J. S. Communicating emotion: Linking affective prosody and word meaning. *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 1017–1030 (2008).
6. Kjelgaard, M. M. & Speer, S. R. Prosodic Facilitation and Interference in the Resolution of Temporary Syntactic Closure Ambiguity. *J. Mem. Lang.* **40**, 153–194 (1999).
7. Speer, S., Warren, P. & Schafer, A. Intonation and sentence processing. in *Proc. 15th Internat. Cong. Phonet. Sci.* 95–105 (2003).
8. Fine, J., Bartolucci, G., Ginsberg, G. & Szatmari, P. The Use of Intonation to Communicate in Pervasive Developmental Disorders. *J. Child Psychol. Psychiatry* **32**, 771–782 (1991).
9. Thurber, C. & Tager-Flusberg, H. Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *J. Autism Dev. Disord.* **23**, 309–322 (1993).
10. Fosnot, S. M. & Jun, S. Prosodic characteristics in children with stuttering or autism during reading and imitation. in *Proc. 14th Internat. Cong. Phon. Sci.*, 1925–1928 (University of California Berkeley, CA, 1999).
11. Shriberg, L. D. Speech and Prosody Characteristics of Adolescents and Adults With High-Functioning Autism and Asperger Syndrome. *J. Speech Lang. Hear. Res.* **44**, 1097–1115 (2001).
12. Diehl, J. J., Bennetto, L., Watson, D., Gunlogson, C. & McDonough, J. Resolving ambiguity: A psycholinguistic approach to understanding prosody processing in high-functioning autism. *Brain Lang.* **106**, 144–152 (2008).
13. Chevallier, C., Noveck, L., Happé, F. & Wilson, D. From acoustics to grammar: Perceiving and interpreting grammatical prosody in adolescents with Asperger Syndrome. *Res. Autism Spectr. Disord.* **3**, 502–516 (2009).
14. Globerson, E., Amir, N., Kishon-Rabin, L. & Golan, O. Prosody Recognition in Adults With High-Functioning Autism Spectrum Disorders: From Psychoacoustics to Cognition. *Autism Res.* **8**, 153–163 (2015).
15. Paul, R., Augustyn, A., Klin, A. & Volkmar, F. R. Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *J. Autism Dev. Disord.* **35**, 205–220 (2005).
16. Wang, A., Dapretto, M., Hariri, A., Sigman, M. & Bookheimer, S. Y. Processing affective and linguistic prosody in autism: An fMRI study. *NeuroImage* **13**, 621 (2001).
17. Hesling, I. *et al.* The Integration of Prosodic Speech in High Functioning Autism: A Preliminary fMRI Study. *PLOS ONE* **5**, e11571 (2010).
18. Eigsti, I.-M., Schuh, J., Mencl, E., Schultz, R. T. & Paul, R. The neural underpinnings of prosody in autism. *Child Neuropsychol.* **18**, 600–617 (2012).
19. Minagawa-Kawai, Y. *et al.* Cerebral laterality for phonemic and prosodic cue decoding in children with autism. *NeuroReport* **20**, 1219–1224 (2009).
20. Molnar-Szakacs, I. & Heaton, P. Music: a unique window into the world of autism. *Ann. N. Y. Acad. Sci.* **1252**, 318–324 (2012).
21. Heaton, P., Pring, L. & Hermelin, B. & others. Musical processing in high functioning children with autism. *Ann. N. Y. Acad. Sci.* **930**, 443–444 (2001).
22. Bonnel, A. *et al.* Enhanced Pitch Sensitivity in Individuals with Autism: A Signal Detection Analysis. *J. Cogn. Neurosci.* **15**, 226–235 (2003).
23. Lai, G., Pantazatos, S. P., Schneider, H. & Hirsch, J. Neural systems for speech and song in autism. *Brain* **135**, 961–975 (2012).
24. Jiang, J., Liu, F., Wan, X. & Jiang, C. Perception of Melodic Contour and Intonation in Autism Spectrum Disorder: Evidence From Mandarin Speakers. *J. Autism Dev. Disord.* **45**, 2067–2075 (2015).
25. Allen, R., Davis, R. & Hill, E. The Effects of Autism and Alexithymia on Physiological and Verbal Responsiveness to Music. *J. Autism Dev. Disord.* **43**, 432–444 (2013).
26. Heaton, P. Assessing musical skills in autistic children who are not savants. *Phil. Trans.: Biol. Sci.* **364**(1522), 1443–1447 (2009).
27. Peretz, I., Vuvan, D., Lagrois, M.-É. & Armony, J. L. Neural overlap in processing music and speech. *Phil. Trans. of the Royal Soc. B, Biol. Sci.* **370**, (2015).
28. Koelsch, S. Toward a neural basis of music perception – a review and updated model. *Audit. Cogn. Neurosci.* **2**, 110 (2011).
29. Friederici, A. D. The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cog. Sci.* **16**, 262–268 (2012).
30. Belyk, M. & Brown, S. Perception of affective and linguistic prosody: an ALE meta-analysis of neuroimaging studies. *Soc. Cogn. Affect. Neurosci.* **9**, 1395–1403 (2014).
31. Just, M. A., Cherkassky, V. L., Keller, T. A. & Minshew, N. J. Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. *Brain* **127**, 1811–1821 (2004).
32. Courchesne, E. & Pierce, K. Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. *Curr. Opin. Neurobiol.* **15**, 225–230 (2005).
33. Just, M. A., Keller, T. A., Malave, V. L., Kana, R. K. & Varma, S. Autism as a neural systems disorder: A theory of frontal-posterior underconnectivity. *Neurosci. Biobehav. Rev.* **36**, 1292–1313 (2012).
34. Frith, U. & Happé, F. Autism: beyond 'theory of mind'. *Cognition* **50**, 115–132 (1994).
35. Happé, F. & Frith, U. The Weak Coherence Account: Detail-focused Cognitive Style in Autism Spectrum Disorders. *J. Autism Dev. Disord.* **36**, 5–25 (2006).
36. Mottron, L., Dawson, M., Soulières, I., Hubert, B. & Burack, J. Enhanced Perceptual Functioning in Autism: An Update, and Eight Principles of Autistic Perception. *J. of Autism and Dev. Disord.* **36**, 27–43 (2006).
37. Zatorre, R. J. & Baum, S. R. Musical melody and speech intonation: singing a different tune. *PLoS Biol.* **10**, e1001372 (2012).
38. Ding, N. *et al.* Temporal modulations in speech and music. *Neurosci. and Biobeh. Rev.* <https://doi.org/10.1016/j.neubiorev.2017.02.011> (2017).
39. Steinhauer, K., Alter, K. & Friederici, A. D. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat. Neurosci.* **2**, 191–196 (1999).
40. Pannekamp, A., Weber, C. & Friederici, A. D. Prosodic processing at the sentence level in infants. *NeuroReport* **17**, 675–678 (2006).
41. Männel, C. & Friederici, A. D. Pauses and Intonational Phrasing: ERP Studies in 5-month-old German Infants and Adults. *J. Cogn. Neurosci.* **21**, 1988–2006 (2009).
42. Itzhak, I., Pauker, E., Drury, J. E., Baum, S. R. & Steinhauer, K. Event-related potentials show online influence of lexical biases on prosodic processing. *NeuroReport* **21**, 8–13 (2010).
43. Pauker, E., Itzhak, I., Baum, S. R. & Steinhauer, K. Effects of Cooperating and Conflicting Prosody in Spoken English Garden Path Sentences: ERP Evidence for the Boundary Deletion Hypothesis. *J. Cogn. Neurosci.* **23**, 2731–2751 (2011).
44. Steinhauer, K. & Friederici, A. D. Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *J. Psycholinguist. Res.* **30**, 267–295 (2001).
45. Steinhauer, K. Electrophysiological correlates of prosody and punctuation. *Brain Lang.* **86**, 142–164 (2003).
46. Knösche, T. R. *et al.* Perception of phrase structure in music. *Hum. Brain Mapp.* **24**, 259–273 (2005).
47. Neuhaus, C., Knösche, T. R. & Friederici, A. D. Effects of Musical Expertise and Boundary Markers on Phrase Perception in Music. *J. Cogn. Neurosci.* **18**, 472–493 (2006).
48. Nan, Y., Knösche, T. R. & Friederici, A. D. Non-musicians' perception of phrase boundaries in music: A cross-cultural ERP study. *Biol. Psychol.* **82**, 70–81 (2009).
49. Silva, S. *et al.* Musical phrase boundaries, wrap-up and the closure positive shift. *Brain Res.* **1585**, 99–107 (2014).

50. Glushko, A., Steinhauer, K., DePriest, J. & Koelsch, S. Neurophysiological Correlates of Musical and Prosodic Phrasing: Shared Processing Mechanisms and Effects of Musical Expertise. *PLOS ONE* **11**, e0155300 (2016).
51. Lord, C. *et al.* The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *J. Autism Dev. Disord.* **30**, 205–223 (2000).
52. Lord, C., Rutter, M. & Couteur, A. L. Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* **24**, 659–685 (1994).
53. [WHO]. World Health Organization. The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines. (Geneva: World Health Organization, 1992).
54. [APA]. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.) (2000).
55. Steinhauer, K., Abada, S. H., Pauker, E., Itzhak, I. & Baum, S. R. Prosody–syntax interactions in aging: Event-related potentials reveal dissociations between on-line and off-line measures. *Neurosci. Lett.* **472**, 133–138 (2010).
56. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *J. Autism Dev. Disord.* **31**, 5–17 (2001).
57. Freitag, C. M. *et al.* Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ) - die Kurzversion AQ-k. *Z. Für Klin. Psychol. Psychother.* **36**, 280–289 (2007).
58. Lehrl, S., Triebig, G. & Fischer, B. Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurol. Scand.* **91**, 335–345 (1995).
59. Horn, W. *Leistungsprüfungssystem, L-P-S: Handanweisung für die Durchführung, Auswertung und Interpretation.* [A performance testing system: Manual for administration, scoring, and interpretation.]. (Verlag Hogrefe, 1962).
60. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
61. Pannekamp, A., Toepel, U., Alter, K., Hahne, A. & Friederici, A. D. Prosody-driven Sentence Processing: An Event-related Brain Potential Study. *J. Cogn. Neurosci.* **17**, 407–421 (2005).
62. Hwang, H. & Steinhauer, K. Phrase Length Matters: The Interplay between Implicit Prosody and Syntax in Korean ‘Garden Path’ Sentences. *J. Cogn. Neurosci.* **23**, 3555–3575 (2011).
63. Braeutigam, S., Swithenby, S. J. & Bailey, A. J. Contextual integration the unusual way: a magnetoencephalographic study of responses to semantic violation in individuals with autism spectrum disorders. *Eur. J. Neurosci.* **27**, 1026–1036 (2008).
64. Koolen, S., Vissers, C. T. W. M., Egger, J. I. M. & Verhoeven, L. Can monitoring in language comprehension in Autism Spectrum Disorder be modulated? Evidence from event-related potentials. *Biol. Psychol.* **94**, 354–368 (2013).
65. Koolen, S., Vissers, C. T. W. M., Egger, J. I. M. & Verhoeven, L. Monitoring in language perception in high-functioning adults with autism spectrum disorder: Evidence from event-related potentials. *Clin. Neurophysiol.* **125**, 108–123 (2014).
66. Istók, E., Friberg, A., Huottilainen, M. & Tervaniemi, M. Expressive Timing Facilitates the Neural Processing of Phrase Boundaries in Music: Evidence from Event-Related Potentials. *PLOS ONE* **8**, e55150 (2013).
67. Hoshi-Shiba, R., Furukawa, K. & Okanoya, K. Neural correlates of expectation of musical termination structure or cadence. *NeuroReport* **25**, 743–748 (2014).
68. Steinhauer, K., Nickels, S., Saini, A. K. S. & Duncan, C. Domain-specific or shared mechanisms of prosodic phrasing in speech and music? New evidence from event-related potentials [abstract]. *J. Cogn. Neurosci. Suppl.* **G-73**, 196 (2009).
69. Drury, J. E., Baum, S. R., Valeriotte, H. & Steinhauer, K. Punctuation and Implicit Prosody in Silent Reading: An ERP Study Investigating English Garden-Path Sentences. *Front. in Psychol.* **7**, 1–12 (2016).
70. Lerdahl, F. & Jackendoff, R. *A Generative Theory of Tonal Music.* (MIT Press, 1985).
71. Merrill, J. *et al.* Perception of words and pitch patterns in song and speech. *Front. in Psychol.* **3**, 76 (2012).
72. Zatorre, R. J., Belin, P. & Penhune, V. B. Structure and function of auditory cortex: music and speech. *Trends in Cog. Sci.* **6**, 37–46 (2002).

Acknowledgements

We would like to thank Sebastian Jentschke for early contributions to this project, Ed Golob for comments on earlier versions of the manuscript, three anonymous reviewers, and all of our participants, without whom this research would not have been possible. We would also like to acknowledge support by the German Research Foundation and the OpenAccess Publication Funds of the Freie Universität Berlin.

Author Contributions

J.D. and S.K. designed the research, J.D. and A.G. collected data, J.D., A.G., and K.S. analyzed the data, and all authors contributed in writing the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-14538-y>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017