

# Neural Responses to Novel and Existing Words in Children with Autism Spectrum and Developmental Language Disorders

Victoria Catherine Penrose Knowland (✉ [victoria.knowland@york.ac.uk](mailto:victoria.knowland@york.ac.uk))

University of York

**Daniel H Baker**

University of York

**Gareth Gaskell**

University of York

**Elaine van Rijn**

University of York

**Sarah A Walker**

University of York

**Courtenay Norbury**

UCL

**Lisa-Marie Henderson**

University of York

---

## Research

**Keywords:** ERP, N400, phonology, ASD, DLD

**Posted Date:** June 26th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-37892/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Journal of Cognition on January 27th, 2022. See the published version at <https://doi.org/10.5334/joc.204>.

# Abstract

**Background** The formation of new phonological representations is key to establishing items in the mental lexicon. Phonological forms become more stable with repetition, time and sleep. This study measured changes in ERP responses to novel and known words within and between two sessions 24 hours apart. We hypothesised that modulation of the auditory N400 would mark changes in lexical processing in children with and without communication difficulties, as novel forms were encoded and consolidated.

**Method** Children listened passively to known and novel words on two consecutive days, while ERPs were recorded using dry electrodes. Data were analysed from 20 typically developing children, 18 children with an autism spectrum disorder (ASD) and 13 children with developmental language disorder (DLD).

**Results** We saw reduced activity overall in our DLD group, reflecting activation in a sparse lexical network, and reduced activity across groups on the second day indicating lexical pre-activation. The DLD group showed the smallest inter-day difference, perhaps signalling reduced overnight maintenance of task-based knowledge. Finally, an effect of lexicality (increased amplitude response to nonwords relative to words) was seen only in the ASD group.

**Conclusion** Data support the dynamic nature of the lexical system, in which neural responses to words and nonwords cause lasting effects on brains' ability to predict incoming stimuli over a 24-hour period, but with these dynamics varying in children with neurodevelopmental disorders.

## 1.0 Introduction

To learn a new word, it is necessary for a novel phonological form to become familiar through exposure. This process is thought to be constrained by perceptual accuracy, the quality of phonological storage and existing phonological representations, as well as the establishment of links to new and existing motor speech programmes (see Gathercole, 2006 for a review). The representation of new phonological forms changes over time: while recognition is often possible after a single hearing, accurate retrieval usually takes repeated exposure, and integration with existing lexical knowledge can be supported by sleep (Dumay & Gaskell, 2007). This prolonged process of acquiring new phonological forms is accounted for theoretically by a Complementary Learning Systems approach to word learning (Davis & Gaskell, 2009), which proposes that new phonological representations are hippocampally mediated, with sleep playing an active role in supporting consolidation of long-term neocortical representations, as supported by fMRI evidence in both adults (Davis & Gaskell, 2009) and children (Landi et al., 2018).

In this study we consider the time-course of early changes in phonological representation and how this might relate to language development in children who differ in their abilities to learn and consolidate words. Using the event-related potential (ERP) technique, we consider change in brain responses to new and established phonological forms over testing sessions, but also change from one day to the next to assess the impact of a consolidation window containing sleep.

# 1.1 Repetition suppression in auditory ERP

One way to measure early changes in neural representation is via repetition suppression: diminishing neurophysiological responses given the repeated presentation of identical or closely related stimuli. Repetition suppression is thought to result from increased processing efficiency (Grill-Spector et al., 2006). It has been presented as an epiphenomenon of predictive coding (see Aukstulewicz & Friston, 2016), whereby it represents an increasing ability to make precise predictions about upcoming sensory events: the closer the match between predicted and experienced events, the smaller the neural response. Here we aimed to measure repetition suppression over the course of each session and from one day to the next.

Repetition suppression is evident in the N400, a negative-going ERP component maximal over central sites 300–500 ms after stimulus onset. Work pioneered by Kutas and Hillyard (see Kutas & Federmeier, 2011), suggests that the N400 represents the neural response to meaningful or potentially meaningful stimuli presented in the auditory or visual domain, and reflects lexical-semantic activation. The amplitude of the N400 is taken as an index of the difficulty of retrieving stored conceptual knowledge through the activation of possible lexical items, with a stronger response seen when the retrieval of lexical information is more taxing. Greater amplitudes are seen in response to phonotactically legal pseudowords than to known words (Helenius et al., 2009), to words and pseudowords with denser lexical neighbourhoods (Laszlo & Federmeier, 2011), and to sentence-final words with low cloze probabilities (see Kutas & Federmeier, 2011). In all these cases (all within session), greater N400 responses are related to increased lexical activation when a phonological form cannot be so easily or predictably mapped to lexical items. Repetition suppression of the N400 (Cheng, Schafer & Riddell, 2014; Marinkovic et al., 2003), suggests that as auditory word forms are repeatedly encountered, phonological storage allows better prediction of that form and less activation of possible lexical matches is required before a match is found. As such, the extent of suppression can be used as a marker for changes in the accessibility of the phonological representation of a previously encountered item.

## 1.2 Lexical activation in atypical populations

Developmental language disorder (DLD) is a neurodevelopmental disorder characterised by a difficulty in the establishment and use of age-typical language skills. Initial evidence suggests that early encounters with novel word forms result in atypical encoding in children with DLD. In a repetition suppression paradigm looking at the modulation of the N400m (the magnetoencephalographic equivalent of the N400) (Helenius et al., 2014), repetition suppression was shown to be absent in response to the second of two presentations of a new word in the left hemisphere of children with DLD relative to typically developing (TD) children. Children with DLD also show well-established behavioural deficits in the ability to repeat novel pseudowords (non-word repetition), particularly given words with three or more syllables (first shown by Gathercole & Baddeley, 1990) or high phonological complexity (Gallon, Harris & van der Lely, 2007). Non-word repetition is thought to mimic early encounters with new words before semantic representations are established, with the quality of early representation within the phonological store

being the foundation for subsequent building of a lexical item (Brown & Hulme, 1996). Notably however, when tested on the same items an hour later, children with DLD show off-line maintenance of new word forms equivalent to TD controls (Bishop, Barry & Hardiman, 2012), albeit performance is lower at both test points; suggesting initial encoding of phonological forms may be more difficult for these children than longer term maintenance of new phonology.

By contrast, verbally-able children on the autism spectrum show enhanced sensitivity to novel phonological forms immediately after exposure relative to typically developing vocabulary-matched peers (Henderson et al., 2014), and are better at matching novel phonological forms to referents (Norbury, Griffiths & Nation, 2010). These children also show non-word repetition abilities in-line with TD children (Williams, Payne & Marshall, 2012). However, lexical representations are not thought to be typical in children with autism spectrum disorders (ASD). Electrophysiological responses to highly familiar lexical-semantic mismatch (e.g., a picture of a FROG paired with the word BRUSH) are weak in verbally-able children with ASD (McCleery et al., 2010). This attenuated neural response may result from a loss of lexical (Henderson et al., 2014) and semantic (Fletcher et al., 2019; Norbury et al., 2010) knowledge about new words over time, compared to TD children. Such semantic memory consolidation difficulties in autism have been linked to atypicalities in sleep parameters (Fletcher et al., 2019). Therefore, while DLD represents a case where early encoding is an issue, but longer-term maintenance may not be, ASD could represent the reverse. These groups therefore provide a natural experiment for testing hypotheses about the factors that are important for long-term memory formation of new words.

## **1.3 The current study**

Here we assess how the time course of encoding and consolidating novel phonological forms differs in TD children compared to those with ASD and DLD. We assess the maturation of the N400 response after 24 hours to examine whether electrophysiological responses to the same items change not only within each session but also across sessions separated by a night of sleep.

It is hypothesised that the TD group should show reduced N400 amplitude (less negative responses) with each presentation of stimuli, especially novel words, and that responses on the second day of testing should be reduced relative to the first, particularly for the novel words as they became more familiar, with responses to known words relatively more stable over days (Davis & Gaskell, 2009). It is hypothesised that the encoding of novel phonological forms within sessions will be less evident in children with DLD compared to TD peers, indexed by reduced repetition suppression of the N400 in response to pseudowords within sessions, but that consolidation between sessions will be comparable across these groups. By contrast, children on the autism spectrum will show more rapid encoding of novel phonological forms within sessions, indexed by increased repetition suppression in response to repeated pseudowords, but reduced consolidation between sessions.

## **2.0 Method**

### **2.1 Participants**

In total, 58 participants were recruited from the UK through local schools and nationwide specialist schools as well as an existing database of children who have previously completed studies in the lab. Twenty-one participants formed the TD group (age range: 8 years 7 months – 12 years 8 months; see Table 1); they had no known developmental disorders, no siblings with disorders, and did not score below 10th percentile on any standardised test in our cognitive battery (described below). For the DLD group, 16 children (age range = 7 years 11–12 years, 3 months) were recruited; these children had a diagnosis of Developmental Language Disorder (or equivalent) and were receiving specialist intervention and/or they scored below 10th percentile on at least two standardised tests of language. Children were recruited to the ASD group if they either had a diagnosis on the autism spectrum, or were awaiting diagnosis, and scored within the ‘very likely’ range on the GARS3. In total, 21 children were recruited to this group (age range = 8 years 10 months-13 years 0 months), though two did not complete testing on both days and were excluded leaving 19 participants. Parents gave informed written consent for their children to take part.

Table 1

Participant characteristics- means (and standard deviations). Age (in months); Male:Female ratio; BPVS-3 standard score; CELF-UK3, Sentence Repetition standard score; BAS-3 Word Definitions, Backward Digit Span and Matrices standard scores; RAN Digits standard score.

	<b>n</b>	<b>Age</b>	<b>M:F</b>	<b>BPVS</b>	<b>S. Rep</b>	<b>Word Defs.</b>	<b>B. Digit Span</b>	<b>Matrices</b>	<b>RAN Digits</b>
<b>TD</b>	21	128.0 (15.2)	15:6	110.0 (8.7)	111.0 (12.9)	109.4 (13.2)	106.6 (18.0)	111.5 (16.1)	107.4 (11.4)
<b>ASD</b>	19	132.2 (15.8)	15:4	102.6 (20.3)	97.1 (20.6)	102.3 (19.5)	101.1 (14.7)	102.7 (20.3)	90.8 (24.8)
<b>DLD</b>	16	123.1 (17.3)	7:9	76.2 (17.5)	70.6 (15.5)	73.0 (19.4)	78.1 (13.6)	77.5 (8.8)	79.7 (13.7)
<b>All</b>	56	128.0 (16.2)	37:19	97.8 (19.9)	94.7 (23.2)	96.6 (22.9)	96.6 (19.6)	98.8 (21.2)	93.8 (20.7)

Children were not invited to participate in the study if English was not their first and dominant language, if their parents or teachers reported a hearing deficit or if they were known to experience generalised epileptic seizures. We also required that children had enough oral language to be able to give informed verbal assent. Two children were known to be taking melatonin to support sleep behaviour at the time of the study, one from the DLD group and one from the ASD group; children were not asked to refrain from their medication for the study. Children were thanked for their time with a £20 Amazon voucher.

## 2.2 Protocol

Each participant completed two sessions (Day 1 and Day 2) approximately 24 hours apart (mean = 23hrs 27 min, SD = 61 min). Participants were seen at the University of York when that was convenient for families, or they were visited at home or in school by one of the authors (VK). Testing locations were always as quiet and calm as possible. At the end of the session on Day 2 participants completed a four-alternative-choice receptive vocabulary check to make sure that all the real words presented during the ERP task were familiar, the target was accompanied by a phonological foil, a semantic foil and a different item from the stimulus set (e.g., apricot, apron, strawberry, pyjamas). Sample-wide mean proportion correct was 0.986 (SD = 0.022), with no difference between groups ( $F(2,58) = 0.774, p = 0.4661$ ).

A cognitive battery was administered to each participant (see Table 1), comprising the following standardised assessments: British Picture Vocabulary Scale, 3rd Edition (BPVS-3; Dunn et al., 2009); The Matrices, Word Definitions, and Backward Digit Recall subscales from the British Ability Scales, 3rd Edition (BAS3; Elliott & Smith, 2011); Recalling Sentences subscale from the Clinical Evaluation of Language Fundamentals; Rapid Automatic Naming (RAN- Digits and Letters) from the Comprehensive Test of Phonological Processing, 2nd Edition (CTOPP2; Wagner et al., 2013). Most participants completed the battery during one of their ERP sessions, while for others some of the battery was completed during a visit prior to the experimental sessions.

The parents of all children were asked to complete a series of questionnaires: The Children's Sleep Habits Questionnaire (CSHQ, Owens, Spirito & McGuinn, 2000); The Children's Communication Checklist, 2nd Edition (CCC-2; Bishop, 2003), and the Gilliam Autism Rating Scale, 3rd Edition (GARS3, Gilliam, 2013). The study was granted ethical approval by the ethics committee for the Department of Psychology at the University of York. Figure 1 shows parent-

## **2.3 Stimuli and paradigms**

### **2.3.1 ERP task**

A mixed factorial design was adopted, with one between-subjects factor: Group (TD, DLD, ASD); and three within-subjects factors: Lexicality (Words, Pseudowords), Presentation (1, 2, 3), and Day (1, 2). Auditory stimuli for the ERP task were 50 real words and 50 pseudowords. For the Word stimuli, 50 concrete nouns were selected (see Supplementary Materials Table SM1 for a list of stimuli) based on being 3–4 syllables in length and having an age-of-acquisition of less than 8.0 years (Kuperman, Stadthagen-Gonzalez & Brysbaert, 2012). There were a few exceptions to the age-of-acquisition rule, deemed acceptable due to perceived changes in usage over time (e.g., computer) or expected differences in UK/US usage (e.g., caravan, badminton).

Pseudowords were created from the 50 Word stimuli (e.g., ALLIBODO derived from ALLIGATOR). This approach was adopted to minimise acoustic and phonotactic differences between the two conditions, while maintaining the lexical-semantic distinction. Pseudowords were onset-matched to their Word counterparts but diverged at the uniqueness point (the point in a spoken word at which it becomes uniquely identifiable). After the uniqueness point no Pseudoword had a phoneme in the same place as in

its counterpart Word, though the two were matched for syllable and phoneme number, and stress, and the stimulus sets did not differ statistically with respect to phonotactic probability. For each Word, uniqueness point was determined using the Celex database taking intonation and age of acquisition into account, so only words which children aged ten years and below (according to Kuperman et al., 2012) were considered in the establishment of uniqueness point. All Word stimuli had at least two phonemes before the uniqueness point and at least two after. See Table SM2 for characteristics of the Word stimuli.

Each stimulus was presented three times, plus 32 catch trials, resulting in a total of 332 trials for each participant on each day (a maximum of 50 trials per cell). The three presentations of each item within a block were not sequential, but occurred with 5–10 intervening items, as demonstrated in Fig. 2. Inter-stimulus-interval was randomly jittered between 2,200-3,200 ms, and stimulus length varied between 626 and 1,289 ms (mean = 948 ms). The experiment was split into four blocks, with each block lasting approximately four minutes, and stimuli were delivered over headphones.

The order of the blocks was randomised across all sessions but the order of stimuli within each block was kept consistent. Participants were not asked to respond to the stimuli, however, they were required to listen for spoken animal noises ('woof', 'moo', 'quack' and 'oink'), and to press a button when they heard one of these noises. Each animal noise was presented twice in each block. These catch trials were included to maintain and monitor auditory attention. Children were not encouraged to press as fast as they could to avoid excessive motoric preparedness responses, but data were analysed to show any group differences in complete lapses, where catch trials were not responded to before the subsequent stimulus was presented. The paradigm lasted approximately 30 minutes on each day.

### **2.3.2 ERP recording**

ERPs were recorded using a portable g.USBamp amplifier, using a sampling frequency of 1200 Hz, with g.SAHARAsys dry active electrodes (not requiring gel). Despite high impedance values compared to lab-based systems, the g.SAHARA system has been shown to record anticipated ERP responses, and is considered among the best of dry electrode systems (Grummett et al., 2015; Radüntz, 2018). Notably though, as far as we are aware, all reliability and validity work has so far been conducted in adult samples (see Lau-Zhu, Lau & McLoughlin, 2019 for a discussion of the benefits and challenges of using mobile EEG equipment with neurodevelopmental populations). Eight channels were used, four along the midline at FPz, Fz, Cz, and Pz, plus two lateral pairs at T7, T8 and C3, C4, anchored in an appropriately sized cap for each child. The reference electrode was placed on the left earlobe and the ground electrode was placed on the forehead. The experimental paradigm and electroencephalographic (EEG) recording were programmed in MATLAB (The MathWorks, 2014). EEG recording for each trial was triggered each time a trial was initiated, with a 687 ms period before the stimulus began; after which stimulus onset responses were recorded for 2,000 ms. A 200 ms baseline was used before stimulus onset.

## **3.0 Results**

### **3.1 Data processing**

Artefact rejection was performed by excluding any trial where peak amplitude exceeded  $\pm 100 \mu\text{V}$  relative to the 200 ms pre-stimulus baseline period. Children were then excluded from analysis if they had fewer than 30 remaining trials on at least one day. On these grounds, the data for 5 children were removed (3 DLD, 1 ASD, 1 TD), leaving 13, 18 and 20 children in these groups respectively. Data points were then rejected if they fell outside  $\pm 2\text{SD}$  of the grand mean. The number of missing trials for remaining participants varied across groups. Of a possible 600 trials in total for each remaining participant, 67.9% of trials were maintained for the TD group, 72.6% of the ASD group and only 50.1% of the DLD group. It is likely that much of the loss in the DLD group can be attributed to movement artefacts with loss of attention, as indicated by the catch trial data. For the TD group, 2.10% of catch trials were lapses, for the ASD group 4.69%, and for the DLD group this rose to 13.57%, resulting in a significant group difference in lapses ( $\chi^2 = 78.026$ ,  $df = 2$ ,  $p < 0.001$ ).

## 3.2 Mixed effects model

Data were analysed in R (R Core Team, 2017) using a multilevel modelling approach, such that ERP trials are not aggregated but entered as individual observations. This approach allowed us to include individual differences in within-participant variance in our models (see Vossen, van Breukelen, Hermens, van Os & Lousberg, 2011 for an assessment of the benefits of using multilevel modelling with ERP data). Models were built with 'lme4' (Bates et al., 2015) with plots made using 'ggplot2' (Wickham, 2016). The package 'Buildmer' (Voeten, 2020) was used for automatic model selection using a forward model selection procedure to establish a parsimonious fixed effects structure, including random intercepts.

The dependent variable was determined after visual inspection of grand mean averages for each electrode to determine the topography and timing of the N400 response in our sample. A linear mixed effects model is therefore presented with average amplitude over 400–500 ms post stimulus onset at four electrodes (Fz, Cz, C3 & C4) as the dependent variable. The following factors were entered as fixed effects into a maximal model: Day (simple coding: Day 1 (-0.5), Day 2 (0.5)), Lexicality (simple coding: Words (-0.5), Pseudowords (0.5)), Presentation (forward difference coding: 1vs2 (1 (2/3), 2(-1/3), 3(-1/3), 2v3 (1/3, 1/3, -2/3), Group (simple coding: TDvASD (TD(-1/3), ASD(2/3), DLD(-1/3), TDvDLD (TD(-1/3), ASD(-1/3), DLD(2/3)), and Electrode (deviation coding with levels Cz,Fz,C3,C4: (-1,0,0,1), (-1,0,1,0), (-1,1,0,0)). Subject and Uniqueness Point were added as random effects. Uniqueness points varied between 127 ms and 527 ms after stimulus onset (mean = 313.7 ms, SD = 92.8); 40 Words and 39 Pseudowords had uniqueness points before 400 ms. These words had a UP on average 120 ms before 400 ms. Visual inspection of Q-Q plots confirmed no substantial deviation from heterogeneity of variance of the residuals for each model.

### 3.2.1 Main effects

The fixed effect of Presentation was maintained in the model as a result of interactions with Day and Lexicality. However, contrary to our hypotheses, no main effect of Presentation emerged (see Table 2). Therefore, the following results consider the global effects of each factor on ERP amplitude rather than the modulation of these effects from repetition suppression, as intended.

Table 2

Fixed and random effects for model at 400–500 ms post stimulus onset. Model formed from 85,283 observations: 51 participants across 2 sessions. Significance levels: \*\*\* at  $p < 0.001$ , \*\*  $p < 0.01$ , \* $p < 0.05$ .

	<i>Fixed effects</i>					<i>Random effects</i>		
	b	SE	95% CI		t	p	Subject	UP
			Lower	Upper			SD	SD
(Intercept)	<b>-5.23</b>	<b>0.32</b>	<b>-5.86</b>	<b>-4.60</b>	<b>-16.187</b>	< 0.001***	2.06	0.95
Day	1.82	0.17	1.49	2.16	10.648	< 0.001***		
Group1(TDvSD)	1.29	0.69	-0.07	2.65	1.854	0.070		
Group2(TDvDLD)	1.89	0.77	0.39	3.40	2.462	0.017*		
Electrode1	0.06	0.14	-0.20	0.33	0.463	0.643		
Electrode2	0.18	0.14	-0.09	0.45	1.334	0.182		
Electrode3	-0.88	0.14	-1.15	-0.61	-6.333	< 0.001***		
Lexicality	-0.55	0.24	-1.02	-0.07	-2.231	0.027*		
Presentation1(1v2)	0.12	0.20	-0.27	0.50	0.592	0.554		
Presentation2(2v3)	0.15	0.19	-0.23	0.53	0.760	0.447		
Day1vGroup1	0.41	0.36	-0.30	1.12	1.136	0.256		
Day1vGroup2	-1.20	0.44	-2.07	-0.33	-2.702	0.007**		
Group1vLexicality	-1.35	0.36	-2.05	-0.65	-3.770	< 0.001***		
Group2vLexicality	-0.29	0.43	-1.14	0.56	-0.672	0.501		
Day vPresentation1	-0.98	0.39	-1.75	-0.21	-2.501	0.012*		
Day vPresentation2	0.59	0.39	-0.17	1.35	1.525	0.127		
Lexicality vPresentation1	0.50	0.39	-0.27	1.27	1.275	0.202		
Lexicality vPresentation2	-0.98	0.39	-1.74	-0.22	-2.517	0.012*		

Main effects of Group emerged between the TD and DLD groups, with the DLD group showing lower amplitude activity (TD mean =  $-6.37 \mu\text{V}$ , se = 0.13, DLD mean =  $-4.33 \mu\text{V}$ , se = 0.18). Likewise, the ASD group tended to show lower amplitude activity than the TD group (see Fig. 4), although the difference did

not quite reach significance (ASD mean =  $-5.02 \mu\text{V}$ , se = 0.13). A main effect of Day was evident, with response amplitude lower on Day 2 (Day 1 mean =  $-6.45 \mu\text{V}$ , se = 0.11, Day 2 mean =  $-4.37 \mu\text{V}$ , se = 0.11). A main effect of Lexicality was seen, with Pseudowords evoking a higher amplitude (more negative) response (Words mean =  $-5.22 \mu\text{V}$ , se = 0.11, Pseudowords mean =  $-5.63 \mu\text{V}$ , se = 0.11). Finally, a main effect of electrode was found, with a significant contrast between C4 and the grand mean (Cz mean =  $-4.80 \mu\text{V}$ , se = 0.17; Fz mean =  $-6.30 \mu\text{V}$ , se = 0.16; C3 mean =  $-5.21 \mu\text{V}$ , se = 0.16; C4 mean =  $-5.38 \mu\text{V}$ , se = 0.15), although Electrode did not interact with any other factors.

### 3.2.3 Interactions

The main effect of Lexicality was influenced by Group and Presentation. These interactions were explored post-hoc by contrasting levels for each Group using 'emmeans' (Lenth, 2019), with Tukey correction for multiple comparisons.

An interaction between Group and Lexicality, emerged between the TD and ASD groups, with Words evoking a less negative (lower amplitude) response in the ASD group compared to the TD group (TD Words, mean =  $-6.46 \mu\text{V}$ , se = 0.18, ASD Words: mean =  $-4.43 \mu\text{V}$ , se = 0.18), while responses to pseudowords were equivalent between the groups (TD Pseudowords, mean =  $-6.28 \mu\text{V}$ , se = 0.18, ASD Pseudowords: mean =  $-5.59 \mu\text{V}$ , se = 0.18). The contrast between Word and Pseudoword conditions only reached significance for the ASD group ( $b = 1.35$ , se = 0.31,  $z = 4.310$ ,  $p < 0.001$ ), with pseudowords evoking a higher amplitude (more negative) response for the ASD but not the TD group (see Fig. 4a).

An interaction between Lexicality and Presentation emerged. The second presentation of word stimuli evoked a lower amplitude (less negative) response than the first, and a higher amplitude response than the third, while the opposite pattern is seen for pseudowords. However, no contrasts emerge as significant.

In addition to showing a robust main effect, Day was found to interact separately with both Group and Presentation. The interaction with Group was driven by the TD and ASD groups showing a greater change between sessions than compared to the DLD group (see Fig. 4, right panel), though the Day contrast was significant for all groups (TD:  $b = -2.09$ , se = 0.25,  $z = -8.240$ ,  $p < 0.001$ ; ASD:  $b = -2.49$ , se = 0.26,  $z = -9.719$ ,  $p < 0.001$ ; DLD:  $b = -0.89$ , se = 0.37,  $z = -2.419$ ,  $p = 0.0155$ ). Finally, the second presentation of stimuli (relative to the first) evoked a larger amplitude (more negative) response on Day 1 (presentation 1 mean =  $-6.12 \mu\text{V}$ , se = 0.20, presentation 2 mean =  $-6.70 \mu\text{V}$ , se = 0.19), but a smaller response on Day 2 (presentation 1 mean =  $-4.52 \mu\text{V}$ , se = 0.20, presentation 2 mean =  $-4.08 \mu\text{V}$ , se = 0.20), although no contrasts were significant. Interactions which did not result in significant contrasts are plotted in Supplementary Materials (Figures SM1&2).

### 3.2.4 Random effects

While both Subject and UP contribute to the model, UP explains very little of the error term, with UP-averaged responses varying from the overall mean with a standard deviation of just 0.95, compared to Subject-averaged responses which vary from the overall mean with a standard deviation of 2.06. This

suggests that the effects we see here relate to a broad period and are not influenced substantially by the uniqueness point of items.

## 4.0 Discussion

This study aimed to assess changes in neural response to novel and known words within and between two ERP sessions spaced 24 hours apart. We asked: does the encoding and consolidation of novel phonological forms differ between typically developing (TD) children and those on the autism spectrum (ASD) or those with developmental language disorder (DLD)? The data suggest that neural responses to words and nonwords evolve over a 24-hour period, and crucially, that such changes (as well as responses overall) are diminished in children coming to the task with varying lexical networks.

Concerning the observed negative-going activity at 400–500 ms post stimulus onset, our clearest result was a robust main effect of day, with responses reducing in amplitude and becoming less negative on the second day of testing (Fig. 4b). Reduced negativity at this latency is typically taken as indicative of reduced lexical search in response to meaningful or potentially meaningful stimuli (Helenius et al., 2009, Laszlo & Federmeier, 2011 and see Kutas & Federmeier, 2011). We expected to see this change specifically in response to novel words as they became more familiar, but not to known words, whose representations should be accessible from the first presentation (Davis & Gaskell, 2009). However, as no interaction between day and condition emerged, the main effect of day here could be interpreted as a reduction in prediction error for both real and pseudoword representations, with a degree of pre-activation evident through the second session. We could speculate that the lack of pseudoword-specific day effects is partially a consequence of the small number of exposures to novel stimuli here relative to previous studies that have found effects of sleep on behavioural responses to pseudowords in children (Brown et al., 2012(12 exposures); Henderson et al., 2012 (18 exposures)). Indeed, differences in exposure level has been shown to influence degree of lexical consolidation in adults (Sobczak & Gaskell, 2019). Three exposures may therefore not be adequate to encourage active engagement with novel items.

Notably, our DLD group showed a smaller difference between days than the TD or ASD groups, suggesting reduced pre-activation on the second day in this group. Previous work has hinted at group differences in the consolidation of linguistic material overnight in adults with developmental language issues. In Earle, Landi & Myers (2018) those with DLD, in contrast to language-typical peers, did not show consolidation effects in tasks requiring the identification and discrimination of new phonetic contrasts. In addition, adults with a history of language issues have been shown to exhibit an increasing deficit in the recollection of novel word forms over the course of a week compared to typically developing peers (McGregor et al., 2013). The role of time and sleep in the consolidation of new phonological and lexical-semantic information in children with developmental language issues needs further exploration to establish the neurophysiological basis for the impaired establishment of novel linguistic forms seen in this population.

Typically, when presented with both real and pseudowords, adult auditory N400 responses show increased negativity (greater amplitude responses) to the pseudowords (Bakker et al., 2015; Bermúdez-Margaretto et al., 2015; Cheng et al., 2014); which is interpreted as indicating increased lexical search in response to phoneme strings that do not match established representations. Here, we do see a main effect of condition, with pseudowords evoking a more negative response than do words. Previous work has shown lexicality effects by at least 100 ms after the point a word can be identified (Gagnepain, Henson & Davis, 2012), and given this, a strong effect of lexicality was expected here even though our real and pseudowords were matched up to the point of uniqueness, meaning that the N400 response was elicited soon after the conditions could be dissociated. Somewhat surprisingly, uniqueness point contributed little to explaining error variance, suggesting that our results apply to stimuli with a range of uniqueness points, perhaps because our analysis considered mean amplitude over the epoch rather than maximal amplitude.

Interestingly, the lexicality effect seen here was driven by the ASD group (see Fig. 4a) rather than the typically developing children. The absence of lexicality effects in the TD group may be because the pseudoword stimuli were derived from the real word stimuli, such that the novel word forms weakly activated existing lexical and semantic representations in the typically developing children, as has previously been seen with written stimuli (Deacon et al., 2004). Some evidence exists to support the idea that lexical-semantic activation may be weaker in individuals with ASD than TD controls, with reduced semantic priming being shown in children with ASD despite age-typical language (Kamio et al., 2007). In addition, children on the autism spectrum have been shown to be more focused on the phonological form of novel words compared to TD peers (Henderson et al., 2014; Norbury et al., 2010). The current data suggest that future work should look to establish the conditions under which N400 lexicality effects are invoked over childhood, as the lexicon expands in size and complexity.

We saw robust effects of group, with the DLD group, showing a less negative response overall compared to the TD group. This group difference may reflect reduced lexical search in children with DLD on account of a sparser lexical network and reduced semantic links. Differences in the breadth and depth of the word knowledge in the DLD group compared to peers with TD or ASD were evident on standardised tests of receptive and expressive vocabulary. So the depth of understanding and semantic context of the real words used as stimuli here are likely to have been poorer in the DLD group even though basic receptive understanding (as captured by the vocabulary check) was equivalent across groups. Importantly, there were also differences in attention between the DLD group and the other two groups, with the DLD group missing many more auditory catch trials. Attention and executive control are known weaknesses in children with DLD (Henry, Messer & Nash, 2011); as such, difficulties with sustained attention are a potential confound when interpreting the contrast between TD and DLD responses. Nonetheless, whether the weak response to verbal material in the DLD group is taken as a measure of semantic integrity or auditory attention, it suggests that auditory processing of incidental verbal material does not elicit the strength of lexical activation as it does in peers with TD or ASD.

Contrary to our expectations, no main effects of repetition were observed here. The absence of clear repetition suppression effects may be a result of our paradigm design. We asked children to respond to catch trials that were unrelated to our stimuli. This was intended to maximise attentional capacity in children and minimise motor preparedness ERP responses. However, as repetition suppression is attenuated or eliminated for unattended stimuli in the visual domain (Eger et al., 2004; Jiang, Summerfield & Egner, 2013; Laszlo, Stites & Ferdermeier, 2012) and for later auditory ERP components (Hsu, Hämäläinen & Waszak, 2014), it is likely that our manipulation may have had the inadvertent effect of reducing attention to the experimental stimuli.

## **4.1 Conclusions, limitations and future work**

The conclusions that we have drawn here should be evaluated within the strengths but also the limitations of the current work. In future replication and extension studies children should be required to attend to the target stimuli in order to invoke repetition suppression, which can then be used as a marker for change in the representation of new lexical items more clearly. The attention issues we observed in the DLD group were significant, and clearly affected the quality of our data set from that group. In general, the use of the dry electrode EEG system, while it had enormous benefits in terms of flexibility of testing, did result in lower signal-to-noise ratio than would be expected in a laboratory setting (Grummett et al., 2015; Radüntz, 2018), with movement artefacts being a substantial issue.

The current study attempted to demonstrate changes in neural response to known and novel phonological forms over the course of two test sessions, 24 hours apart in children with typically developing language systems, those with autism spectrum disorders and those with developmental language disorder. We saw weak (less negative) responses in children with DLD, with reduced change overnight compared to peers with TD or ASD. This suggests that children with DLD showed less pre-activation of lexical representations in response to words encountered the previous day, such that lexical access was facilitated to a smaller degree. We saw a lexicality effect emerge in children with ASD, but not their typically developing peers, which we speculate is due to word-like stimuli activating existing representations in children with TD. This finding could be supported by work to show that children with ASD are less likely to lexicalise heard legal nonwords. Overall, our data support the idea of change in phonological representations over time and suggest that that change may be diminished in children with sparser lexical networks.

## **Declarations**

## **Ethics approval and consent to participate:**

study approval was granted by the departmental ethics committee for the Department of Psychology at the University of York, UK. The parents of all participants gave written, informed consent for their children to participate in the study, and all children gave verbal assent.

## Consent for publication:

N/A

## Availability of data and materials:

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

## Competing interests:

None.

## Funding:

This work was supported by the Economic and Social Research Council [grant number ES/N009924/1, 2015].

## Authors' contributions:

VK designed the study, collected all data, analysed results and wrote the manuscript. DB coded the paradigm and extracted the data. GG designed the study and edited the manuscript. EvR edited the manuscript. SW designed stimuli and edited the manuscript. CN designed the study. LH designed the study and edited the manuscript. All authors read and approved the final manuscript.

## Acknowledgements:

The authors would like to thank the children and families who took part in this work, and the ESRC for funding it.

## References

1. Auzstulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex*, *80*: 125-140.
2. Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Tracking lexical consolidation in ERPs: lexical and semantic-priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, *79* (A): 33-41.

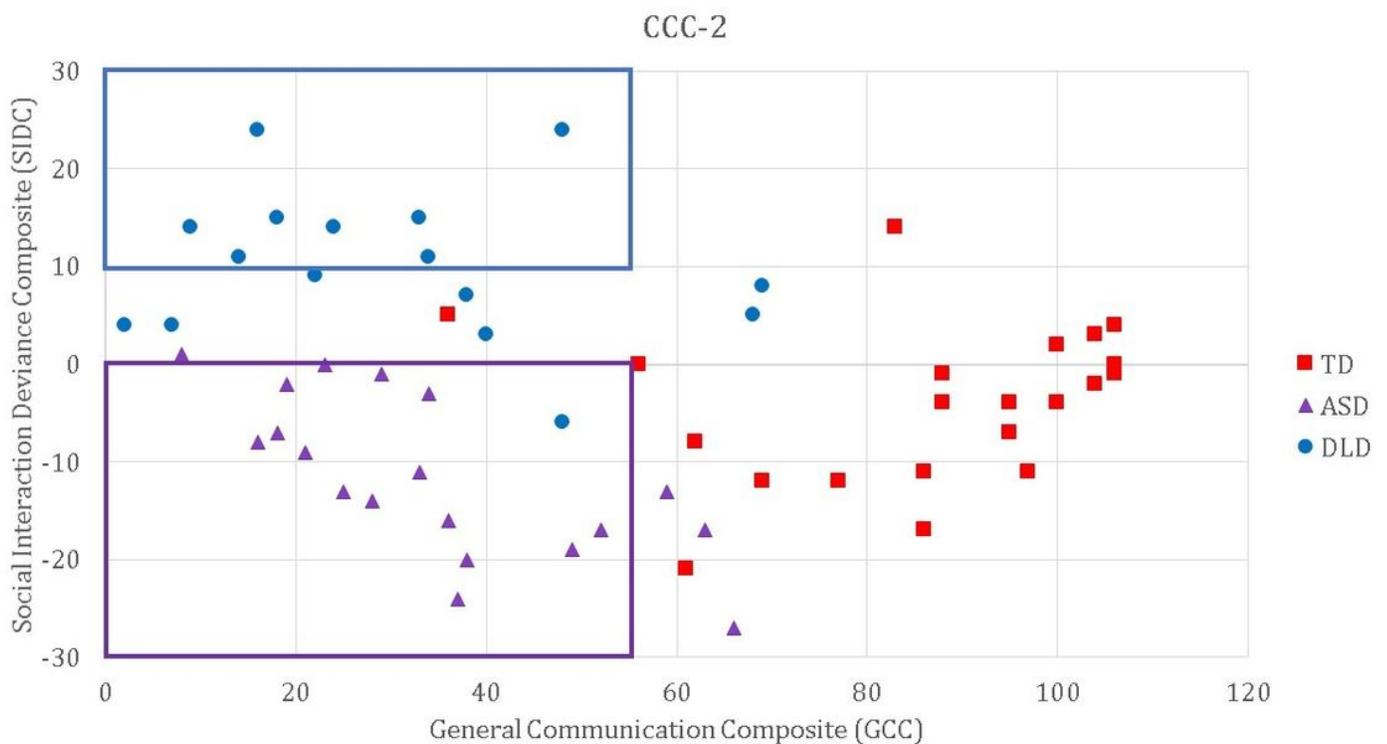
3. Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.
4. Bermúdez-Margaretto, B., Beltrán, D., Domínguez, A., & Cuetos, F. (2015). Repeated exposure to “meaningless” pseudowords modulates LPC, but not N(FN)400. *Brain Topography*, 28, 838-851.
5. Bishop, D. V. M. (2003). *The Children’s Communication Checklist, 2<sup>nd</sup> Edition*. London: Pearson.
6. Bishop, D. V. M., Barry, J. G., & Hardiman, M. J. (2012). Delayed retention of new word-forms is better in children than adults regardless of language ability: A factorial two-way study. *PLoS one*, 7(5): e37326.
7. Brown G. D. A., & Hulme C. (1996). *Nonword repetition, STM and word age of acquisition: A computation model*. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 129–148). Hove: Psychology Press.
8. Brown, H., Weighall, A., Henderson, L. M., and Gaskell, M. G. (2012). Enhanced recognition and recall of new words in 7- and 12-year-olds following a period of offline consolidation. *Journal of Experimental Child Psychology*, 112: 56–72. doi: 10.1016/j.jecp.2011.11.010
9. Cheng, X., Schafer, G., Riddell, P. M. (2014). Immediate auditory repetition of words and nonwords: an ERP study of lexical and sublexical processing. *PLoS ONE* 9(3): e91988. doi:10.1371/journal.pone.0091988
10. Davis, M., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of Royal Society of London B: Biological Sciences*, 364(1536):3773-3800. doi: 10.1098/rstb.2009.0111
11. Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. *Psychophysiology*, 41(1): 60-74.
12. Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, 18 (1): 35-39.
13. Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). *The British Picture Vocabulary Scale III – 3<sup>rd</sup> Edition*. London: GL Assessment.
14. Earle, F. S., Landi, N., & Myers, E. B. (2018). Adults with Specific Language Impairment fail to consolidate speech sounds during sleep. *Neuroscience Letters*, 666, 58-63.
15. Eger, E., Henson, R.N.A., Driver, J., & Dolan, R.J. (2004). BOLD repetition decreases in object-responsive ventral visual areas depend on spatial attention. *Journal of Neuroscience*, 24: 1241-1247. doi:10.1152/jn.00206.2004
16. Elliott, C. D., & Smith, P. (2011). *The British Ability Scales, 3<sup>rd</sup> Edition*. London: GL Assessment.
17. Fletcher, F., Knowland, V.C.P., Walker, S.A., Gaskell, M.G., Norbury, C., & Henderson, L-M. (2019). *Developmental Science*, 23 (3): e12906. doi:10.1111/desc.12906
18. Gagnepain, P., Henson, R., & Davis, M. H. (2012). Temporal predictive codes for spoken words in auditory cortex. *Current Biology* 22, 615–621. doi 10.1016/j.cub.2012.02.015

19. Gathercole, S. (2006). Nonword repetition and word learning: the nature of the relationship. *Applied Linguistics*, 27(6) doi: 10.1017/S0142716406060383
20. Gathercole, S., & Baddeley, A. (1990). Phonological memory deficits in language disordered children: is there a causal connection? *Journal of Memory and Language*, 29(3): 336-360. doi: 10.1016/0749-596X(90)90004-J
21. Gallon, N., Harris, J., & van der Lely, H. K. J. (2007). Nonword repetition: an investigation of phonological complexity in children with grammatical-SLI. *Clinical Linguistics and Phonetics*, 21, 435-455.
22. Gilliam, J.E. (2013). *Gilliam Autism Rating Scale-3<sup>rd</sup> Edition*. Pearson: London, UK.
23. Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Science*, 10: 14-23. doi:10.1016/j.tics.2005.11.006
24. Grummett, T. S., Leibbrandt, R. E., Lewis, T. W., DeLosAngeles, D., Powers, D. M.W., Willoughby, J. O., Pope, K. J., & Fitzgibbon, S. P. (2015). Measurement of neural signals from inexpensive, wireless and dry EEG systems. *Physiological Measurement*, 36: 1469–1484 doi:10.1088/0967-3334/36/7/1469
25. Helenius, P., Parviainen, T., Paetau, R., & Salmelin, R. (2009). Neural processing of spoken words in specific language impairment and dyslexia. *Brain*, awp134.
26. Helenius, P., Sionen, P., Parviainen, T., Isoaho, P., Hannus, S., Kauppila, T., Salmelin, R., & Isotalo, L. (2014). Abnormal functioning of the left temporal lobe in language-impaired children. *Brain & Language*, 130: 11-18.
27. Henderson, L. M., Weighall, A. R., Brown, H., & Gaskell, G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science*, 15 (5), 674-687.
28. Henry, L., Messer, D., & Nash, G. (2011). Executive control in children with specific language impairment. *Journal of Child Psychology and Psychiatry*, 53 (1): 37-45. doi: 10.1111/j.1469-7610.2011.02430.x
29. Henderson, L.M., Powell, A., Gaskell, M.G., & Norbury, C. (2014). Consolidation of vocabulary in autism spectrum disorder. *Developmental Science*. Early on-line view, 2014 Mar 17. doi:10.1111/desc.12169.
30. Hsu, Y-F., Hämäläinen, J. A., & Waszak, F. (2014). Repetition suppression comprises both attention-independent and attention-dependent processes. *NeuroImage*, 98: 168-175. doi: 10.1016/j.neuroimage.2014.04.084
31. Jiang, J., Summerfield, C., & Egnor, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *The Journal of Neuroscience*, 33 (47): 18438-18447.
32. Kamio, Y., Robins, D., Kelley, E., Swainson, B., & Fein, D. (2007). Atypical lexical/semantic processing in high-functioning autism spectrum disorders without early language delay. *Journal of Autism and Developmental Disorders*, 37:1116–1122 doi: 10.1007/s10803-006-0254-3 123
33. Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44 (4): 978-990.

34. Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62*: 621–647.
35. Landi, N., Malins, J. G., Frost, S. J., Magnuson, J. S., Molfese, P., Ryherd, K., Rueckl, J. G., Mencl, W. E., Pugh, K. R. (2018). Neural representations for newly learned words are modulated by overnight consolidation, reading skill, and age. *Neuropsychologia, 111*: 133-144.  
doi:10.1016/j.neuropsychologia.2018.01.011
36. Lau-Zhu, A., Lau, M. P. H., & McLoughlin, G. (2019). Mobile EEG in research on neurodevelopmental disorders: opportunities and challenges. *Developmental Cognitive Neuroscience, 36*: 100635.  
doi:10.1016/j.dcn.2019.100635
37. Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbour and lexical associate effects. *Psychophysiology, 48*(2):176-86.
38. Laszlo, S., Stites, M., & Federmeier, K. D. (2012). We won't get fooled again: An event-related potential study of task and repetition effects on the semantic processing of items without semantics. *Language and Cognitive Processes, 27* (2): 257-274.
39. Length, R. (2019). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.3.5. <https://CRAN.R-project.org/package=emmeans>
40. Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron, 38*: 487–497.
41. The MathWorks. (2014). *MATLAB Release 2014*, The MathWorks, Inc., Natick, Massachusetts, United States.
42. McCleery, J. P., Ceponiene, R., Burner, K. M., Townsend, J., Kinnear, M., & Schreibman, L. (2010). Neural correlates of verbal and nonverbal semantic integration in children with autism spectrum disorders. *The Journal of Child Psychology and Psychiatry, 51* (3), 277-286.
43. McGregor, K. K., Licandro, U., Arenas, R., Eden, N., Stiles, D., Bean, A., & Walker, E. (2013). Why words are hard for adults with developmental language impairments. *Journal of Speech, Language and Hearing Research, 56* (6), 1845-1856.
44. Norbury, C. F., Griffiths, H., & Nation, K. (2010). Sound before meaning: word learning in autistic disorders. *Neuropsychologia, 48* (14), 4012-4019.
45. Owens, J. A., Spirito, A., & McGuinn, M. (2000). The Children's Sleep Habits Questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children. *Sleep, 23* (8): 1043-1051.
46. R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
47. Radüntz, T. (2018). Signal quality evaluation of emerging EEG devices. *Frontiers in Physiology, 14*:  
[doi.org/10.3389/fphys.2018.00098](https://doi.org/10.3389/fphys.2018.00098)
48. Sobczak & Gaskell, M.G. (2019). Implicit versus explicit mechanisms of vocabulary learning and consolidation. *Journal of Memory and Language, 106*:1-17. [doi:10.1016/j.jml.2019.01.003](https://doi.org/10.1016/j.jml.2019.01.003)

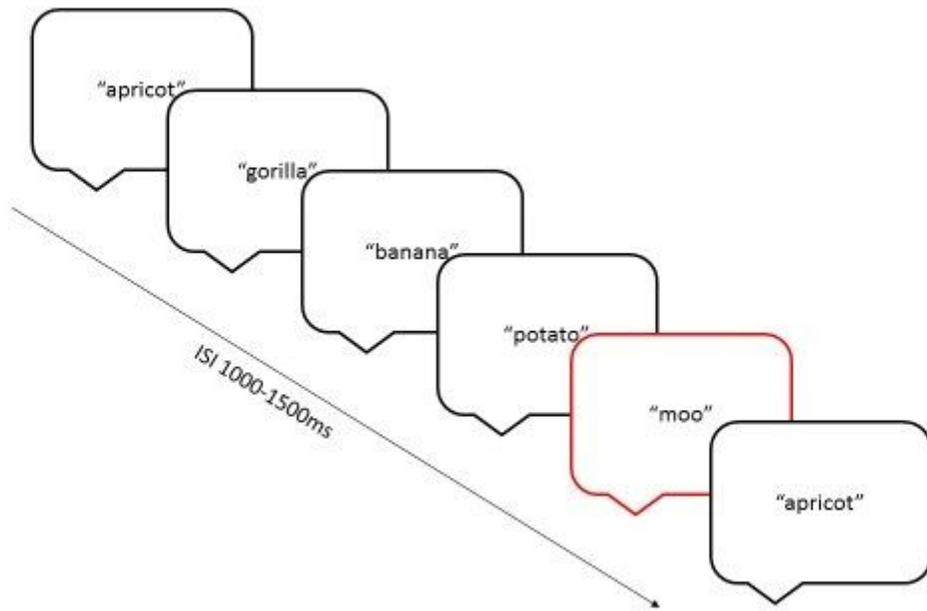
49. Vossen, H., van Breukelen, G., Hermens, H., van Os, J., & Lousberg, R. (2011). More potential in statistical analyses of event-related potentials: a mixed regression approach. *International Journal of Methods in Psychiatric Research*, 20(3): e56– e68.
50. Voeten, C.C. (2020). buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression. R package version 1.5. <https://CRAN.R-project.org/package=buildmer>
51. Wagner, R., Torgesen, J., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive test of phonological processing, 2<sup>nd</sup> Edition* (CTOPP2). Oxford: Pearson.
52. Wickham, H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
53. Williams, D., Payne, H., & Marshall, C. (2012). Non-word repetition impairment in autism and specific language impairment: evidence for distinct underlying cognitive causes. *Journal of Autism and Developmental Disorders*, 43 (2): 404-417.

## Figures



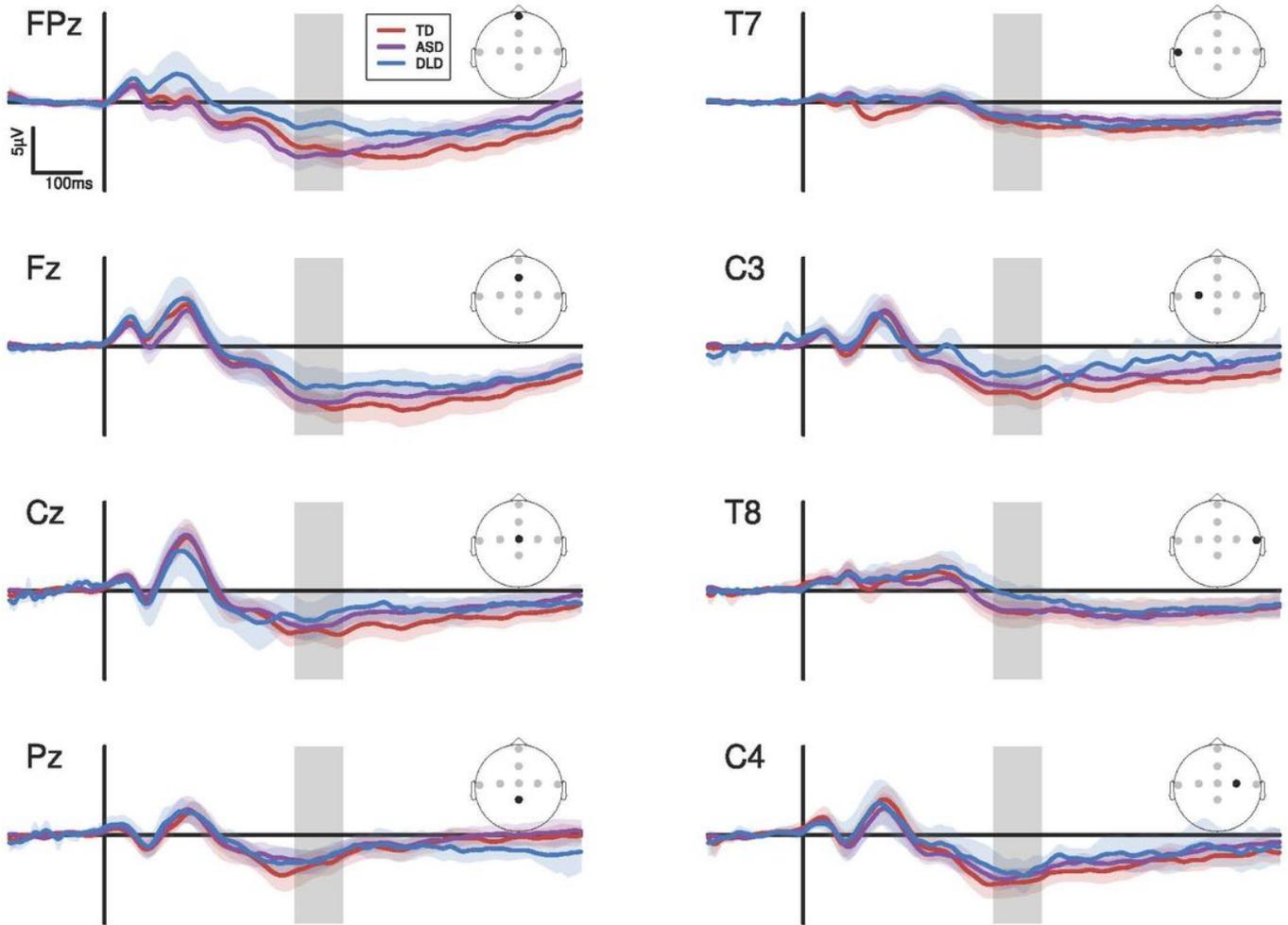
**Figure 1**

Children's Communication Checklist (2nd edition) scores. A GCC score below 55 in conjunction with a negative SIDC score (marked by the purple box) indicates social communication difficulties consistent with autism spectrum conditions; a GCC score below 55 in conjunction with an SIDC of 10 or more (marked by the blue box) is consistent with a structural language difficulty.



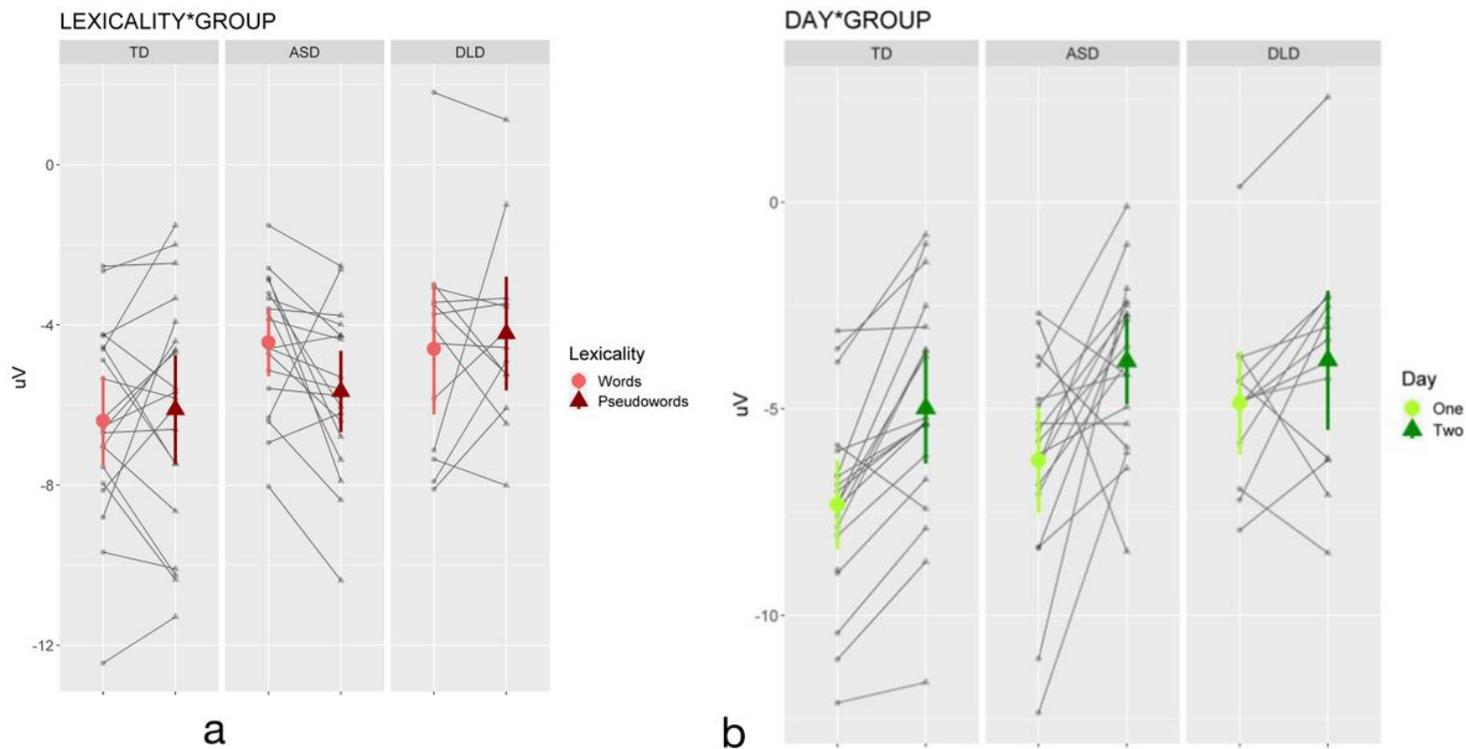
**Figure 2**

Example string of items.



**Figure 3**

Grand average activity at each electrode for each group. Grey rectangles indicate 400-500ms post stimulus onset. Shaded regions surrounding the curves indicate bootstrapped 95% confidence intervals.



**Figure 4**

Interactions between a) Lexicality and Group and b) Day and Group, with mean values given; error bars show 95% CI.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterials.docx](#)