

Title: **Shorter neural adaptation to sounds accounts for dyslexics' abnormal perceptual and reading dynamics**

Authors: Sagi Jaffe-Dax¹, Orr Frenkel², Merav Ahissar^{1,2}

Affiliations: 1. Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 9190401, Israel.

2. Psychology Department, The Hebrew University of Jerusalem, Mt. Scopus Campus, Jerusalem 9190501, Israel.

Abstract

Dyslexia is a prevalent reading disability whose underlying mechanisms are still disputed. We studied the neural mechanisms underlying dyslexia using a simple frequency-discrimination task. Though participants were asked to compare the 2-tones in each trial, implicit memory of previous trials affected their responses. We hypothesized that implicit memory decays faster among dyslexics. We tested this by increasing the temporal intervals between consecutive trials, and measuring the behavioral impact and ERP responses from the auditory cortex. Dyslexics showed a faster decay of implicit memory effects on both measures, with similar time constants. Finally, faster decay also characterized dyslexics' benefits in oral reading rate. It decreased faster as a function of the time interval from the previous reading of the same non-word. We propose that dyslexics' shorter neural adaptation paradoxically accounts for their longer reading times, since it induces noisier and less reliable predictions for both simple and complex stimuli.

Introduction

Dyslexics are diagnosed on the basis of their persistent difficulties in acquiring peer-level reading skills despite adequate education. Their general reasoning skills are within the normal range (or above), but they consistently show difficulties in some language related skills such as verbal working memory (e.g. Torgeson and Goldman, 1977), and phonological manipulations (which typically also load on short term memory; e.g. Landerl et al., 1997). Dyslexics also often have higher thresholds in simple perceptual discrimination tasks (McAnally and Stein, 1996; Witton et al., 1998; Hämäläinen et al., 2013), particularly when administered with serial presentations (Ben-Yehudah and Ahissar, 2004; discussed in Ramus and Ahissar, 2012). In most of these protocols participants can be more successful by taking into account the statistics of previous stimuli (Ahissar et al., 2006; Oganian and Ahissar, 2012).

The putative causes of dyslexics' difficulties on simple serial tasks have been studied in a series of works using 2-tone frequency discrimination. Ahissar et al. (2006) measured the impact of sound regularities on dyslexics' performance. They assessed a well-documented observation (Harris, 1948), that listeners use a repeated reference tone as an “anchor” and improve their performance compared with a no-reference protocol. Dyslexics' benefit from this repetition was smaller than good readers'. A similar deficit was found in dyslexics benefit from repetition of speech sounds. This led to the hypothesis that dyslexics have a deficit in using sound stimuli as perceptual anchors for the formation of sound predictions (Ahissar et al., 2006; Ahissar, 2007; Oganian and Ahissar, 2012).

Raviv et al. (2012) extended the protocol-specific account of benefits from stimulus repetition to a computational model which takes the experiment's statistics into account. This model assumes that listeners implicitly infer the mean (frequency) even when it is not presented

explicitly. The inferred mean of previous stimuli (prior) is combined with the representation of the current stimulus, and forms an integrated percept (posterior). The resulting percept is contracted (biased) towards that mean (the “contraction bias”; Woodrow, 1933; Preuschhof et al., 2010). This bias is advantageous when the observation of the current stimulus is noisy, and hence integration with prior knowledge is likely to improve its accuracy. Indeed, in the general population, “noisier” listeners weigh the prior more than “less noisy” listeners when they integrate it with current representation, resulting in their larger contraction bias. Jaffe-Dax et al. (2015) found that dyslexics’ bias is smaller than controls’ even though they tend to be “noisier” listeners.

The representation of the first tone is contracted towards the prior more than the second tone, due to the noise added to its representation during the inter-stimulus time interval. The contraction of the first tone towards the prior can increase the perceived difference between the two tones in the trial, and hence improve discrimination (*Bias+* trials; e.g., trial t in Fig. 1A); or, decrease the perceived difference between the two tones and disrupt performance (*Bias-*; e.g., trial $t - 1$ in Fig. 1A). The difference in performance between *Bias+* and *Bias-* trials reflects the magnitude of the contraction bias (i.e. context effect). Dyslexics’ underweighting of the prior yields a smaller contraction bias, namely a smaller performance difference between these two types of trials (Jaffe-Dax et al., 2015).

Using this measure of contraction bias, we studied the dynamics of controls’ and dyslexics’ benefits from the statistics of recent sound stimuli in both simple discriminations and in oral reading. We found that dyslexics are similarly affected (biased) by recent stimuli, but less affected by earlier stimuli, as expressed both behaviorally and in the dynamics of the compulsory ERP components (N1 and P2) produced by the auditory cortex. These observations suggest that dyslexics’ automatic integration of previous sounds spans shorter time intervals, and is therefore

noisier and produces less reliable priors. We propose that noisier predictions impede dyslexics' acquisition of expert-level performance in a range of tasks, including reading. These observations also pave the way to pursuing the deficits underlying dyslexia in non-human animals.

Results

Dissociating the contraction to the most recent trial from the contraction to all previous trials

Performance in the 2-tone frequency discrimination task is affected by two basic components (Raviv et al., 2012, 2014). First, the frequency difference between the two tones composing each trial - participants are more accurate when this difference is larger. Second, the context effects; namely, the perceptual bias induced by contracting the first tone towards the estimated mean-frequency in previous trials (Woodrow, 1933; Preuschhof et al., 2010; see Introduction and Fig. 1A-B). This estimated prior can be viewed as a combination of a recent component and a global component. The most recent one is the frequency of the previous trial (Fig. 1B), whose impact on the estimated prior is substantially larger than that of any other previous trial ("recency effect"; Schab and Crowder, 1988; Stewart and Brown, 2004; Fischer and Whitney, 2014). The global prior (illustrated in Fig. 1A) is the mean frequency of all previous trials.

As described above, dyslexics' discrimination behavior is less affected by the statistics of previous trials (Jaffe-Dax et al., 2015). We now aimed to decipher whether dyslexics' deficit encompasses both the recent and the global context effects. To test this, we designed a novel protocol for the 2-tone frequency discrimination task, which allowed us to assess the contribution of recent and earlier trials' on performance separately. Unlike in most sequences where the

recent and more global contexts tend to have the same direction, and are therefore very difficult to dissociate, in the new sequence the recent and the global contexts were not correlated (the constraints that this sequence satisfies are described in the Methods).

The overall accuracy of the two groups in this task did not differ (controls' mean % correct \pm SEM: 75.3 ± 1.6 , dyslexics': 73 ± 1.9 ; $z = 1.3$, *n.s.* Mann-Whitney U test), as predicted given that the use of experiment's statistics was not expected to be significantly beneficial in this stimulus series (replicating Ahissar et al., 2006). We calculated a GLM model with 3 predictors (β s) for each participant ($n = 60$; 30 control and 30 dyslexic participants) by estimating the magnitude of the contribution of each of the following components to participants' responses: 1. Frequency difference in the current trial; 2. Contraction bias of the first tone toward the global mean of previous trials; 3. Contraction bias of the first tone towards the most recent trial. Dyslexics' frequency difference predictor did not differ from controls' (Fig 1C; $z = 1.6$, *n.s.*). Namely, the frequency difference within the trial had a similar impact on the response in the two groups (similar levels of sensitivity). The impact of the context effect of the most recent trial (ITI ≈ 1.5 sec) was also similar in the two groups ($z = 0.05$, *n.s.*). However, dyslexics' bias towards the global mean was significantly smaller than controls' ($z = 2.7$, $p < 0.01$). Indeed, the difference in controls between the contributions of the *Global* (all previous trials except the most recent one) and *Recent* contexts was larger than that found in dyslexics ($z = 2.4$, $p < 0.05$. Mann-Whitney U tests). Importantly, among both groups, the contribution on *Recent* context was significantly above zero (controls: $z = 4.4$, $p < 0.0001$; dyslexics: $z = 4.7$, $p < 0.00001$). Thus floor effect cannot account for this interaction.

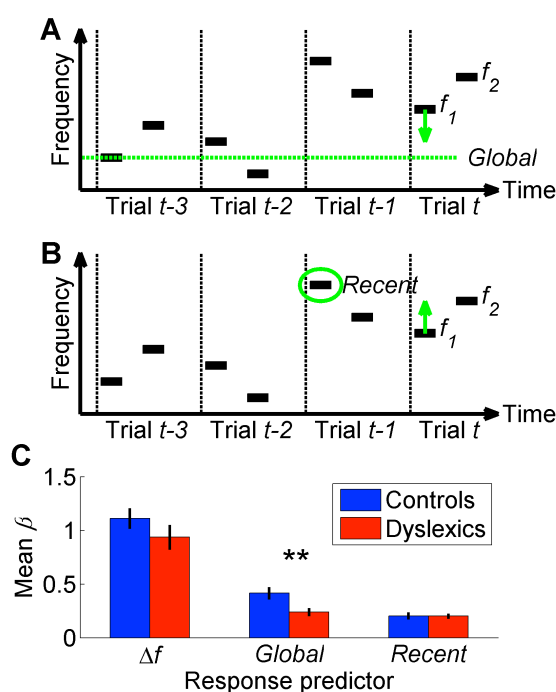


Figure 1. The three parameters that additively determine performance in frequency discrimination: *frequency difference within a trial* (Δf), *Global context* (the difference between the current f_1 and the global mean) and *Recent context* (the difference between the current f_1 and the previous f_1). **A.** A schematic illustration of the *Global bias* - four trials and the direction of the contraction, which contracts the representation of the first tone in the trial towards the global mean. **B.** A schematic illustration of the same four trials and the direction of the *Recent bias*, which pulls the representation of this tone towards the recent frequency. **C.** The estimated contribution (β) of each the three parameters to the overall performance of each group. Dyslexics differed from controls only in the smaller magnitude of the contribution of their *Global* predictor ($p < 0.01$). Filled bars denote the mean β values; controls in blue and dyslexics in red. Error bars denote SEM.

Comparing the behavioral and neural dynamics of context effects

The observation that dyslexics assign less weight to earlier trial statistics than controls can result from either of following two mechanisms. The first is that dyslexics experience larger interference effects, and intervening trials mask each other. The second is that dyslexics' memory trace decays faster in time, even without intervening sounds. To dissociate between these alternatives, we manipulated the time interval between consecutive trials. We reasoned that if dyslexics' memory trace decays faster we might be able to track the neural correlate of this behavioral dynamics.

We administered the 2-tone frequency discrimination task with four different Inter-Trial Intervals (ITIs; i.e. interval between the 2nd tone of the previous trial and the 1st tone of the current trial), in four separate blocks. We chose ITIs of 1.5, 3, 6 and 9 sec (roughly; see Methods), based on previous reports of cortical adaptation duration (N1 and P2 components; Hari et al., 1982; Lu et al., 1992; Sams et al., 1993).

Assessing the time constant of the behavioral context effect

We quantified behavioral performance in terms of sensitivity to the difference between the two tones. As a measure of sensitivity we used d' , which was calculated as the difference between the correct rate (HR) and the incorrect rate (FA), transformed by inverse cumulative standard distribution function (Macmillan and Creelman, 2004): $d' = \Phi^{-1}(HR) - \Phi^{-1}(FA)$. For both groups, performance was better at longer ITIs (mean $d' \pm \text{SEM}$ for the four ITIs was 0.48 ± 0.05 , 0.57 ± 0.05 , 0.71 ± 0.06 and 0.78 ± 0.06 , respectively; $\chi^2 = 54.2$, $p < 10^{-10}$. Friedman test). The slight improvement in d' with longer ITIs is probably related to having “time outs” for responding. These “time-outs” were enforced due to the co-registration of ERP with behavioral performance, and probably introduced a slightly larger stress with short ITIs.

Across all ITIs, controls' ($n = 23$) performance was slightly more accurate than dyslexics' performance ($n = 25$; mean $d' \pm \text{SEM}$ of controls: 0.71 ± 0.07 , dyslexics: 0.56 ± 0.07 ; $z = 2.0$, $p < 0.05$. Mann-Whitney U test). There was no significant interaction between group and ITI (group X ITI; $F_{3,46} = 1.4$, *n.s.* Repeated Measures ANOVA test). Controls' slightly better performance is interesting, particularly since no group difference was found for the same participants in Experiment 1. The group difference was small, and was within the range of the estimated group variability. It may, however, reflect slightly better learning of the task's characteristics by the control participants.

We calculated the behavioral context effect as the difference (in d') between performance on trials for which contraction to the prior was beneficial (*Bias+*) and performance on trials for which this contraction was disruptive (*Bias-*). Previous studies have found that the impact of earlier stimuli decayed exponentially as a function of number of trials which had passed (Raviv et al., 2012) or the time interval (Lu et al., 1992) or both (Fischer and Whitney, 2014). Thus, to quantify the dynamics of this effect, i.e. its magnitude as a function of the time interval from previous trials, we fitted the calculated bias ($\Delta d'$; difference between individual d' in *Bias+* trials and in *Bias-* trials) as a function of the ITI to an exponential decay model: $\Delta d'(t) = \alpha + \beta \exp(-t/\tau)$, where α denotes the estimated $\Delta d'$ at $t \rightarrow \infty$ (asymptotic level); β denotes the difference between the $\Delta d'$ at $t = 0$ and at $t \rightarrow \infty$ (decay magnitude); and τ denotes the time it takes for $\Delta d'$ (at $t = 0$) to decay to $1/e$ (~37%) of its initial value (temporal slope parameter). A small τ indicates fast decay; t is the ITI.

Evaluating the dynamics of the contraction bias as a function of ITI showed that its decay was significantly faster among dyslexics, as illustrated in Fig. 2B bottom (controls' $\tau = 6 \pm 0.9$ sec, dyslexics' $\tau = 2.9 \pm 0.8$ sec; mean $\pm \text{SEM}$; $z = 2.2$, $p < 0.05$). The two other parameters did not

differ between the two groups (α : controls = $0.7 \pm 0.1 d'$, dyslexics: $\alpha = 0.5 \pm 0.1 d'$; $z=1$, *n.s.*;
 β : controls = $2.3 \pm 0.8 d'$, dyslexics = $4 \pm 0.9 d'$; $z = 1.9$, *n.s.* Mann-Whitney U tests), in line
 with our hypothesis.

Assessing the dynamics of the neural trace

We hypothesized that the neural mechanism that mediates the inference of the prior frequency
 is neural adaptation, which is an automatic, stimulus-specific form of memory (Ulanovsky et al.,
 2003). Importantly, it was found that the time constant of context effects in sound discrimination
 is the same as that of cortical adaptation to sounds (Lu et al., 1992). The two prominent ERP
 components that are automatically produced by the auditory cortex are N1 and P2 (Sheehan et
 al., 2005; Mayhew et al., 2010), which peak at 100 and 200ms from stimulus onset, respectively.
 Whether they are produced by a single cortical generator (evolving in time) or by two is still
 debated (Lanting et al., 2013). Both were found to have some sensitivities to stimuli statistics
 (Tremblay et al., 2010; Herrmann et al., 2015). However, in our previous study where we
 assessed sensitivity to frequency prior, only the magnitude of P2 showed such sensitivity. We
 therefore hypothesized that P2 would be directly associated with the contraction bias. Still, given
 the mixed results regarding the distinction between N1 and P2 sensitivities, we analyzed the
 dynamics of both components. We hypothesized that for each group, behavioral context effects
 and cortical adaptation (N1 and/or P2) would have the same time constants and that these time
 constants would be shorter among dyslexics.

In both groups, the magnitude of the P2 responses was the smallest in the block with the
 shortest ITI (1.5 sec), when P2 adaptation was substantial, and increased (recovered) with longer
 ITIs, as adaptation gradually decayed ($\chi^2 = 65.7$, $p < 10^{-13}$. Friedman test). Among dyslexics, P2
 reached its peak magnitude by 6 sec ITI (Fig. 2A, red dot-dashed line), whereas controls' P2 was

larger at 9 than at 6 sec ITIs (Fig. 2A, blue dotted line). To quantify the dynamics of P2 adaptation under active discrimination, we fitted each individual's areas of P2 elicited by the first tone, in each of the 4 ITIs to the same equation of exponential decay. Figure 2B *top* shows the cumulative distribution of the 3 fitted parameters for each of the two groups. It shows that the groups did not differ in the estimated amplitudes of P2 at either maximal or minimal adaptation (α : controls = $396 \pm 47 \mu V \times ms$, dyslexics = $322 \pm 33 \mu V \times ms$; $z = 1.1$, *n.s.*; β : controls = $-696 \pm 99 \mu V \times ms$, dyslexics = $-696 \pm 103 \mu V \times ms$; $z = -0.1$, *n.s.* Mann-Whitney U tests). However, as with behavior, the rate at which adaptation decayed was significantly faster among dyslexics (τ : controls = 6.3 ± 0.8 sec, dyslexics = 2.2 ± 0.8 sec; mean \pm SEM; $z = 3.5$, $p < 0.0005$. Mann-Whitney U test).

As was evident when comparing the time constant of the recovery of P2 (time constant of its adaptation) and the time constant of the behavioral context effect, the two were very similar. This similarity was also found in the dyslexic group. Figure 2C plots the estimated dynamics (the fitted $\alpha + \beta \exp(-t/\tau)$ curves) of P2 adaptation and the behavioral contraction bias in the same graph. For clarity, the dynamics of the ERP component is plotted as decaying to zero, i.e. the fitted curve was subtracted from α (the asymptotic level of maximal P2 magnitude) to depict the decay of adaptation. Dyslexics' behavioral and P2 adaptation curves were very similar. Controls' behavioral and P2 adaptation were also similar and both were significantly slower than dyslexics' (Fig. 2B).

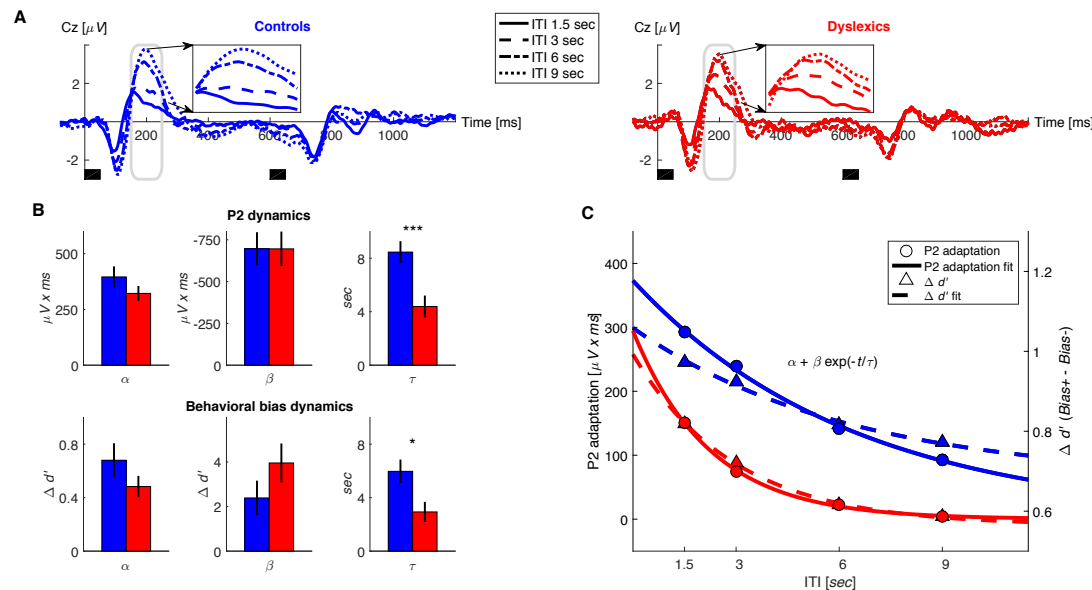


Figure 2. Dyslexics' decay of both neural adaptation and behavioral contraction bias was faster than controls'. **A.** Grand average ERPs, plotted separately for blocks with different ITIs, for controls ($n = 23$; blue, left) and dyslexics ($n = 25$; red, right). Timing of the 2 tones in each trial is denoted by the short black bars under the plots. **Middle insets:** P2 range (denoted in gray) enlarged. Dyslexics' P2 area was similar for 6 and 9 second ITIs, whereas controls' P2 was larger for the 9 second interval. **B.** Groups' mean and SEM of individually fitted estimated parameters of the exponential decay model of P2 recovery from adaptation (**top**) and contraction bias decay (**bottom**). In both, dyslexics differed from controls solely on the estimated rate of decay (P2: $p < 0.0005$; $\Delta d'$: $p < 0.05$). **C.** The decrease in P2 adaptation (solid lines, circles, left scale; estimated magnitude at asymptote minus fitted curve) and the decrease in contraction bias (dashed lines, triangles, right scale; reflecting implicit memory decay) as a function of ITI. Data in markers, fitted exponential decay model in lines. Controls in blue, dyslexics in red.

We further analyzed the dynamics of N1. In this component as well, dyslexics showed a faster recovery from adaptation than controls. As illustrated in Figure 3 A and C, dyslexics' N1 almost fully recovered (regained full magnitude) at an ITI of 3 seconds, whereas controls' N1 at this ITI was still substantially "adapted". Applying the same model, again we found that only the rate of decay differed between the two groups (Fig. 3B; τ : controls = 4.5 ± 0.8 sec, dyslexics = 0.7 ± 0.4 sec; mean \pm SEM; $z = 4.2$, $p < 0.0001$. Mann-Whitney U test), whereas the two other estimated parameters were similar (α : controls = $-131 \pm 20 \mu V \times ms$, dyslexics = $-109 \pm 16 \mu V \times ms$; $z = 0.7$, *n.s.*; β : controls = $497 \pm 120 \mu V \times ms$, dyslexics = $750 \pm 132 \mu V \times ms$; $z = 1.1$, *n.s.* Mann-Whitney U tests). Across groups, N1 adaptation was significantly shorter than that of P2 ($z = 3.7$, $p < 0.0005$. Wilcoxon test). There was no group X ERP component interaction effect for the rate of decay of adaptation ($z = 0.7$, *n.s.* Mann-Whitney U test).

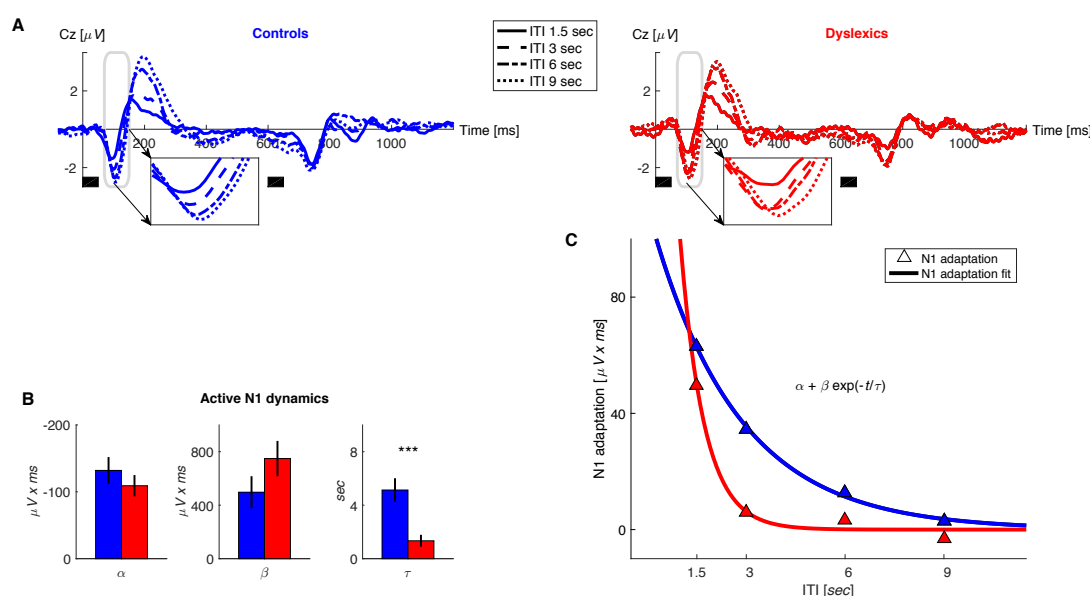


Figure 3. Dyslexics' decay of N1 adaptation was faster than controls'. **A.** Grand average ERPs, plotted separately for blocks with different ITIs, for controls ($n = 23$; blue, left) and dyslexics (n

= 25; red, right). Timing of the 2 tones in each trial is denoted by the short black bars under the plots. **Middle insets:** N1 range (denoted in gray) enlarged. Dyslexics' N1 almost reached its full magnitude at an ITI of 3 sec, whereas controls' N1 reached its near-full magnitude only at an ITI of 6 sec. **B.** Groups' mean and SEM of individually fitted estimated parameters of the exponential decay model of N1 recovery from adaptation. Only τ differed between the two groups ($p < 0.0001$). **C.** The decrease in N1 adaptation (circles) and the fitted exponential decay model (lines) as a function of ITI. Controls in blue, dyslexics in red.

Comparing the dynamics of passive adaptation between controls and dyslexics

After showing that adaptation under active performance decays faster among dyslexics, we further asked whether this group-difference was an automatic characteristic that could also be detected in the absence of a behavioral task. To evaluate the dynamics of dyslexics' passive adaptation, we administered a sequence of pure tones in four blocks, each with a different ISI, while participants watched a silent movie (ISIs: 2, 3.5, 6.5 and 9.5; same TOAs as those used for the active condition, but without the second tone).

To quantify the dynamics of adaptation of P2 and N1, we fitted the P2 and N1 areas, to the exponential decay model described above. We found that controls' P2 recovery was significantly slower than dyslexics'. Importantly, the only significant difference between the two groups was in τ (τ : controls = 4.2 ± 0.8 sec, dyslexics = 2.1 ± 0.7 sec; mean \pm SEM; $z = 3.3$, $p < 0.005$. Mann-Whitney U test; Fig. 4B). The other two parameters did not differ between the two groups (α : controls = $490 \pm 73 \mu V \times ms$, dyslexics = $410 \pm 44 \mu V \times ms$; $z = -0.04$, *n.s.* β : controls = $-750 \pm 112 \mu V \times ms$, dyslexics = $-724 \pm 100 \mu V \times ms$; $z = -0.1$, *n.s.* Mann-Whitney U tests).

N1 adaptation was not significantly shorter among dyslexics (τ : controls = 5 ± 0.9 sec, dyslexics = 2.6 ± 0.9 sec; $z = 1.7$, $p = 0.05$. α : controls = $-132 \pm 20 \mu V \times ms$, dyslexics = $-114 \pm 16 \mu V \times ms$; $z = 0.37$, *n.s.* β : controls = $497 \pm 120 \mu V \times ms$, dyslexics = $615 \pm 132 \mu V \times ms$; $z = 0.02$, *n.s.* mean \pm SEM; Mann-Whitney U tests). In line with previous studies of passive presentation (Hari et al., 1982), we did not find a difference between the time scales of adaptation of P2 and N1 ($z = 1.6$, *n.s.* Wilcoxon test) nor was there a group X ERP component interaction ($z = 1.4$, *n.s.* Mann-Whitney U test).

Figure 4A shows the average ERP responses in the different blocks, for controls ($n = 23$; blue, left) and dyslexics ($n = 25$; red, right), respectively. It focuses on P2, since only P2 showed a significant group difference (insets in Figure 4A). In both groups, P2 responses were the smallest in the block with the shortest ISI (2 sec), when P2 adaptation was still substantial, and increased (recovered) with longer ISIs, as adaptation gradually decayed ($\chi^2 = 60.3$, $p < 10^{-12}$. Friedman test). Among dyslexics, P2 reached its peak magnitude by 6.5 sec ISI (Fig. 4A, red dot-dashed line), whereas controls' P2 was larger at 9.5 than at 6.5 sec ISIs (Fig. 4A, blue dotted line).

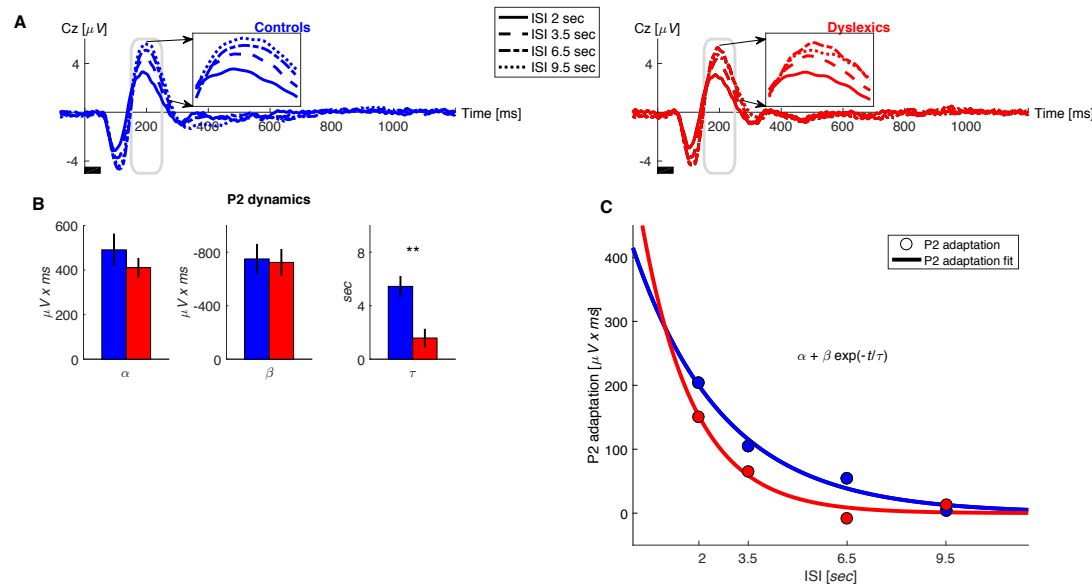


Figure 4. Under passive conditions, dyslexics' P2 adaptation is shorter than controls'. **A.** Grand average ERPs, plotted separately for blocks with different ISIs, for controls (blue, left) and dyslexics (red, right). **Middle insets:** P2 range (denoted in gray) enlarged. Dyslexics' P2 area was largest for the 6.5 second ISI, whereas controls' P2 was largest for the 9.5 second ISI. **B.** Groups' mean and SEM of individually fitted estimated parameters of the exponential decay model of P2 recovery from adaptation. Controls in blue; dyslexics in red. Dyslexics differed from controls only in the estimated rate of decay ($p < 0.005$). **C.** The magnitude of P2 adaptation as a function of time from the previous stimulus (ISI) under passive condition (triangles) and fitted exponential decay model (solid line). Controls in blue; dyslexics in red.

Interestingly, among controls, the estimated duration of adaptation was longer in the active than in the passive condition (P2: $z = 3$, $p < 0.005$. N1: $z = 4.2$, $p < 0.00005$). This was not observed in the dyslexic group (P2: $z = 0.5$, *n.s.* N1: $z = 0.1$, *n.s.* Wilcoxon tests). The interaction (in the estimated time constant of adaptation) between group and condition (passive/active) was

significant for N1 ($z = 2.7$, $p < 0.01$), but not for P2 adaptation ($z = 1.8$, *n.s.* Mann-Whitney U tests). Thus, controls' but not dyslexics' N1 adaptation was prolonged by introducing a behavioral context.

Comparing the dynamics of context effects in reading rate

As shown above, dyslexics benefit less from recent sound statistics. We hypothesized that a similar observation would be found for recent encounters with complex sound stimuli and particularly with novel syllabic combinations. If demonstrated, this correlate would provide support for the relevance of the observations of Experiments 1-2 to the context of reading. To assess the dynamics of the simplest context effect in reading, we administered an oral reading task in which participants were asked to read single simple disyllabic non-words aloud, and measured repetition effects. The non-words were presented on a screen, one at a time, and participants were asked to read them as quickly (and accurately) as they could. Consecutive non-words were presented at an individually paced rate. Non-words were repeated throughout the block with various intervals (i.e. with a different number of other intervening non-words), as schematically illustrated in Figure 5A.

Participants' response times (RTs, measured as vocal response onset) were shorter in the second presentation of the same non-word. However, the magnitude of this improvement declined as the time interval from the previous presentation of the same non-word increased (i.e. there were more intervening words). Overall, dyslexics' RTs were longer than controls' (controls: $n = 29$; mean RT per word \pm SEM = 725 ± 28 ms; dyslexics: $n = 23$; mean RT \pm SEM = 929 ± 47 ms; $z = 3.3$, $p < 0.001$. Mann-Whitney U test) and they were less accurate (controls:

mean % correct \pm SEM = 98.6 ± 0.3 ; dyslexics: 96.0 ± 1.4 ; $z = 3.2$, $p < 0.005$. Mann-Whitney U test), as expected.

Both groups showed large recent context (repetition) effects. Specifically, both groups showed substantially faster reading RTs in the second than in the first encounter with the same non-word, when the words were proximal in time (< 2 sec interval), even though they were never consecutive. The smallest interval between two presentations of the same non-word was with one intervening word. Since word presentation rate was based on reading rate, and dyslexics are slower readers, we equated the inter-word time interval in the two groups by considering only trials in which the second presentation occurred within < 2 sec ITI (offset of the first presentation to the onset of the second presentation of the same non-word). In these trials dyslexics' benefit from repetition was similar to controls' (Fig. 5B; controls: 25 ± 2 % or 170 ± 19 ms, dyslexics: 23 ± 3 % or 177 ± 20 ms; mean benefit \pm SEM; $z = 0.1$, *n.s.* Mann-Whitney U test). However, in trials with larger inter-word intervals, dyslexics' benefits were overall smaller. Importantly, dyslexics' benefits were smaller than controls', even though their starting point had longer RTs, as explained above (mean ITI of 42 sec, max ITI of 163 sec. controls: 12 ± 1 % or 90 ± 10 ms, dyslexics: 6 ± 1 % or 52 ± 8 ms; mean benefit \pm SEM; $z = 3$, $p < 0.005$. Mann-Whitney U test). Indeed, among dyslexics the difference in benefit from repetitions with small (< 2 sec) compared to larger (> 2 sec) intervals was greater than this difference among controls ($z = 2.2$, $p < 0.05$. Mann-Whitney U test).

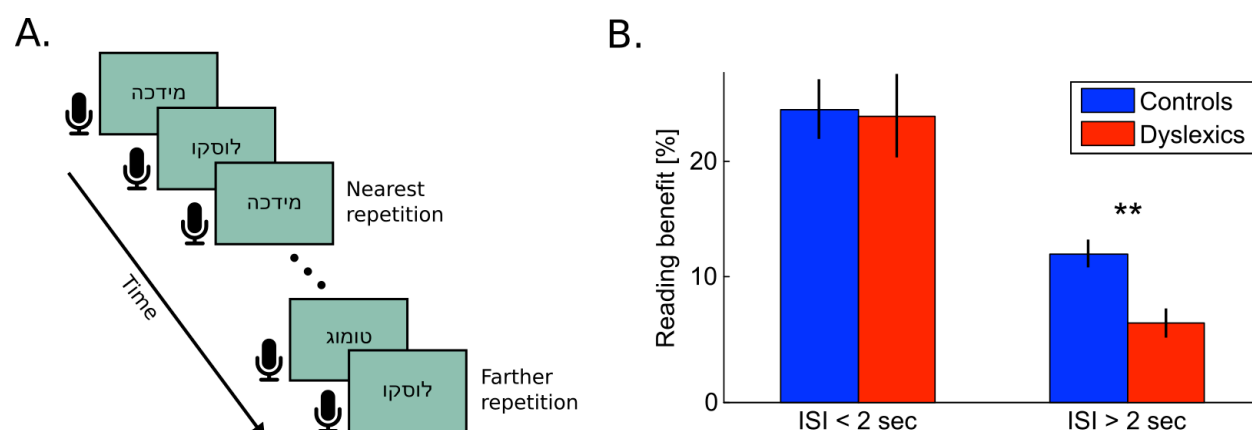


Figure 5. Dyslexics' benefit from a previous exposure to the same non-word decayed faster than controls'. **A.** Schematic illustration of the reading task. Subjects were asked to read the non-words aloud as quickly as possible. Presentation switched to the next word with the subject's voice offset. The closest repetition of the same non-word was with one intervening non-word (i.e. an ITI of < 2 sec). **B.** Benefit in RTs (response times from visual word presentation to vocal onset) as a function of the time interval between the first and second presentation of the same non-word. Improvement was calculated as the difference in RT between the first and second presentation of the same non-word in the block. At very short intervals (< 2 sec) the benefit was similar for both groups. However, this benefit decayed faster (at interval > 2 sec) among dyslexics (red) than among controls (blue; $p < 0.005$). Error bars denote standard error (SEM).

Note that for short inter-word intervals, dyslexics' RT benefits do not differ from controls' in spite of an intervening word. Thus, here too, dyslexics' faster decay is consistent with a time based effect rather than with an enhanced interference effect.

Discussion

We characterized the dynamics of context effects, and found that they differ consistently between dyslexics and good readers. Dyslexics' behavioral benefits declined faster, both in the context of reading and in sound discrimination. In both groups, the behavioral dynamics reflected the dynamics of cortical adaptation, which decayed faster in dyslexics, both when actively performing a task and in passive listening.

Accounting for Dyslexics' "slower processing speed"

We propose that dyslexics' shorter time constant of adaptation reflects a shorter time constant of implicit sound integration. Dyslexics' faster decay decreases the time constant of their integration of sound statistics, and consequently reduces the reliability of their implicitly calculated priors. Due to their noisier priors, attaining the same level of processing reliability requires extraction of more on-line information, which requires more processing time, leading to the stereotypic "slower processing time" in dyslexia.

This account explains dyslexics' often poorer performance on rapidly presented brief auditory stimuli for both simple tones (Tallal, 1980; Temple et al., 2000) and speech (McArthur and Bishop, 2005; Boets et al., 2007). Moreover, it resolves seemingly inconsistent observations in dyslexics' pattern of perceptual difficulties. On the one hand, dyslexics' poor short term memory impairment is amply documented (Beneventi et al., 2010; Banai and Yifat, 2012) and suggests that their performance should improve when memory load is reduced by decreasing task's retention intervals. On the other hand, dyslexics are known to perform poorly with short stimuli presented at brief intervals (Tallal, 1980), suggesting that they should gain from longer inter-stimulus intervals.

We claim that these seemingly contradictory observations result from the same basic impairment. The key feature is the proportion of on-line computations that need be allocated to solve the task. The more predictable the stimuli are for controls, both locally within session and globally in terms of long-term regularities (such as in language) the greater relative slowness (“slow processing speed”) dyslexics are expected to show (consistent with Ahissar et al., 2006). This slowness stems from their poor implicit memory mechanisms; namely, their faster decay of the neural trace that integrates the statistics of recent stimuli.

Dyslexics’ deficit as a process rather than a representational impairment

Dyslexics’ phonological and reading skills reflect the outcome of long-term learning processes which are difficult to track. The use of a simple task of discrimination of pure tones (which are physically simple but very infrequent environmentally) enabled us to observe the dynamics of context effects in a relatively novel situation. It served to dissociate the benefits of very short-term (< 2 sec; the scale attributed to working memory processes; Baddeley, 2010) implicit effects from somewhat longer term effects. Dyslexics’ only benefitted less from prior statistics in the latter case.

Dyslexics’ noisier priors (predictions based on past experiences) may lead to slower long-term learning of characteristic sound regularities, such as those in the native language (Nicolson et al., 2010). It follows that dyslexics should benefit less than controls from a given number of experiences, and that group differences should increase with additional specific experiences. Thus, though both good and poor readers are expected to improve with practice, dyslexics’ benefits per exposure are expected to be smaller. Hence, counter intuitively, dyslexics’ relative difficulties are expected to be greater for highly trained stimuli.

Our observation of reduced repetition effects in dyslexics' reading rate is compatible with imaging observations by Pugh et al. (2008), who compared the impact of word repetition in reading on BOLD activation in posterior reading-related areas, in dyslexics versus controls. They found that although both populations gained from such repetitions, the impact differed between groups. Whereas controls' activation was monotonically reduced with repetitions, dyslexics' activation initially increased. They interpreted this pattern in a manner consistent with ours by suggesting that dyslexics' pattern seemed similar to that expected after a lesser amount of exposure to these words; i.e., slower long-term learning of these (sound) patterns. Interestingly, dyslexics' slower learning rate per event (or exposure) was also recently proposed by Nicolson et al., (2010), who based their argument on slower learning curves in a different context. They found that dyslexics exhibit slower learning of simple motor tasks, which perhaps tap mechanisms partially common to those used in serial comparison tasks.

The complex relationships between mechanisms and sites

We found that dyslexics' P2 component had a shorter adaptation duration under both active and passive conditions. We also found that their N1 had a shorter adaptation duration under active condition when compared to controls. The P2 (which peaks 200ms after tone onset) component was shown to reflect the accumulation of sound statistics (Tremblay et al., 2010; Jaffe-Dax et al., 2015) automatically, without explicit attention (Sheehan et al., 2005). Moreover, it was shown to be abnormal among dyslexics (Bishop and McArthur, 2004; Jaffe-Dax et al., 2015). Indeed, its dynamics matched that of the behavioral context effects measured simultaneously. However, even the earlier N1 component (100ms) had a shorter time constant here among dyslexics under the active condition. Whether these two components originate from one or two cortical

generators is still unclear (e.g. Mayhew et al., 2010; Lanting et al., 2013). Nevertheless, recent studies have found that the magnitude of adaptation of both is sensitive to the statistics that characterize the experiment (e.g. Herrmann et al., 2015). We found a shorter time constant of adaptation in N1 than in P2.

These observations suggest that the dynamics of adaptation in the auditory cortex of dyslexics differs from that of controls. But they do not directly point to the anatomical source of this abnormality, since by 100-200 ms ERP signals reflect the combined contribution of various brain sites. Thus, the abnormally fast decay of N1 and P2 adaptation may reflect a different anatomy of the auditory cortex, as suggested in a recent longitudinal study (Clark et al., 2014), or impaired long-range connectivity between posterior and frontal areas (Boets et al., 2013; Ramus, 2014). Alternatively, it may correspond to impaired top-down control (Díaz et al., 2012) or unreliable bottom-up processing (Hornickel and Kraus, 2013).

Importantly, the involvement of the auditory cortex does not exclude the contribution of sub-cortical regions. Both the cerebellum and the basal ganglia are involved in the process of integrating sound regularities into improved perceptual performance with reduced reliance on on-line working memory processes and increased reliance on specific sound predictions (Daikhin and Ahissar, 2015). Thus, the hypothesis linking dyslexics' reduced rate of automatization (in reading and in other tasks) with impaired usage of cerebellar processes (Nicolson et al., 2010) is also in line with our findings and interpretation.

The attempts to map dyslexics' processing difficulties to several potentially relevant anatomical abnormalities suggest that there may not be a single anatomical source of difficulty, and that the issue of the core deficit underlying dyslexia, should be seen as a process- rather than structure-related. Our observation that the process of dyslexics' neural adaptation is abnormally short

paves the way to studies of reading difficulties in non-humans, which are essential for understanding the complex relationships between the dynamics of cognitive processes and their underlying neural mechanisms.

The prevalence of shorter adaptation in dyslexia

About half of our dyslexic participants showed adaptation time scales that were within the range of controls (Figs. 2-4). This partial overlap between the two groups could have stemmed from the low reliability of our measurements. Another potential source for this variability might be different types of dyslexia among our dyslexic participants, which may relate to different underlying neural mechanisms. Our standard reading tests that were used for inclusion or exclusion of participants might not have been fine-grained enough to define participants along the dimensions that map to different sources of underlying difficulties (Zoccolotti and Friedmann, 2010).

The putative functional relationship between shorter adaptation and long-term effects

We only tracked the short-term effects of dyslexics' faster decay of memory trace. Therefore, we can only speculate about the relationship between these observations and dyslexics' difficulties in acquiring expert-level proficiency in reading. One putative conceptual link may be provided by the Bayesian framework, when the principle of efficient coding is introduced (Wei and Stocker, 2015). This principle predicts that more likely stimuli gradually acquire denser, more reliable representations. The gradual enhancement of the representation of more likely stimuli may be reduced in dyslexia, leading to reduced sensitivity to the specific phonological

and morphological forms that characterize their native language. Assessing this hypothesis experimentally requires long-term tracking of the learning of novel sound statistics.

Materials and Methods

Participants

Sixty native Hebrew speakers (30 dyslexics and 30 good readers), all of whom were students at the Hebrew University [mean age (STD) = 24.2 (5.4) years; 36 females] were recruited for this study. Recruitment was based on ads at the Hebrew University. Monetary compensation for participation was according to standard student rates. The study was approved by the Hebrew University Committee for the Use of Human Subjects in Research. All dyslexic participants had been diagnosed as having a specific reading disability by authorized clinicians. Reading related measures were also assessed in our lab (detailed in Table 1). Participants with more than 2 years of formal musical education were excluded, so that musical training would not be a major contributor to their pitch sensitivity (Micheyl et al., 2006; Parbery-Clark et al., 2011). Participants with poor Block Design scores (lower than a normalized score of 7) were also excluded from the study. All participants filled in a questionnaire regarding any neurological or psychiatric disorders. None of the participants reported any such disorders. None of them had ever participated in a similar auditory experiment in our lab.

Experimental Procedure

Participants were administered 4 sessions on 4 different days:

In session 1 participants were administered a series of cognitive assessments. 30 dyslexics and 30 controls were admitted to this session.

In session 2 participants performed a 2-tone frequency discrimination task with a specially designed sequence of trials (Experiment 1; Fig. 1A-B). The same 30 dyslexics and 30 controls participated in this session.

In session 3, ERPs were recorded both passively and while performing the discrimination task. First, participants watched a silent movie while a series of single tones was presented to them in four blocks of four different Inter Stimulus Intervals (ISI) of 2, 3.5, 6.5 or 9.5 seconds (Experiment 2b), in a random order. Second, participants actively engaged in the 2-tone frequency discrimination task in four blocks with different Inter-Trial-Intervals (ITI – time interval between the second tone in the trial and first tone in the following trial) of 1.4, 2.9, 5.9 and 8.9 seconds (Experiment 2a). The ISI between the two tones in the trial was 600 ms. Thus, in both passive and active conditions, the Stimulus Onset Asynchronies (SOA), i.e. the time intervals between the onset of the first tones in adjacent trials were 2, 3.5, 6.5 and 9.5 seconds. A subgroup of 25 dyslexics and 23 controls participated in this session.

In session 4 participants performed a fast reading task of visually presented single non-words. Voice response was recorded (Experiment 3; Fig 4A). Both rate and accuracy were obtained. 29 dyslexics and 23 controls participated in this session.

All sessions were administered in a sound-attenuated room. Sounds and visual presentations were produced using Matlab (The Mathworks, Inc., Natick, MA). Tones were presented and

voice response was recorded by Psychtoolbox and Psychportaudio (Kleiner et al., 2007) through a Saffire 6 USB audio interface (Focusrite Audio Engineering Ltd., High Wycombe, UK).

Cognitive assessments

General cognitive abilities and phonological skills were assessed using standard tasks:

A. Non-verbal reasoning ability. This was measured with the Block Design, a standard test for assessing visuo-spatial reasoning (WAIS-III; Wechsler, 1997).

B. Short-term verbal memory. This was evaluated with the standard Digit Span task (forward and backward; Hebrew version of WAIS-III; Wechsler, 1997).

C. Phonological decoding and single word reading. Pseudo-word and single word reading were assessed using standard Hebrew lists designed by Deutsch and Bentin (1996).

D. Word pattern recognition. Subjects were presented with 24 pairs, each composed of a word and a pseudo-homophone, and were asked to point to the word in each pair.

E. Fluent reading. Subjects read an academic level passage of 150 words followed by a comprehension question.

F. Phonological awareness was assessed using the Spoonerism task (MacKay, 1970; Möller et al., 2007). Participants heard (Hebrew) word pairs and were asked to switch the first phonemes of the two words and respond vocally (e.g.: /laila tov/, “good night” in Hebrew, should be switched to /taila lov/).

In all phonological and reading tasks both accuracy and rate were scored.

Test	Control (STD) N=30	Dyslexic (STD) N=30	Mann- Whitney z value
Age (years)	25.8 (3.0)	24.3 (3.1)	1.6
General cognitive (scaled)			
Block Design	12.4 (2.9)	12.1 (3.5)	0.3
Digit Span	11.1 (2.8)	7.8 (1.7)	4.7***
Phonological Speed [items/minute]			
Pseudo-word Reading Rate	58.4 (24.4)	32.2 (10.5)	4.4***
Single-word Reading Rate	96.8 (32.5)	68.3 (25.8)	3.3**
Word Pattern Recognition Rate	68.1 (15.2)	39.9 (13.5)	5.6***
Passage Reading Rate	140.4 (23.8)	97.8 (22.3)	5.7***
Spoonerism Rate	10.0 (3.0)	5.8 (3.2)	4.6***
Phonological Accuracy [% correct]			
Pseudo-word Reading Accuracy	89.7 (11.2)	62.4 (18.3)	5.1***
Single-word Reading Accuracy	97.1 (4.3)	87.6 (8.3)	4.8***
Word Pattern Recognition Accuracy	100 (0)	96.2 (6.4)	4.3***
Passage Reading Accuracy	98.6 (1.4)	94.8 (4.5)	4.9***
Spoonerism Accuracy	92.2 (6.8)	77.9 (18)	3.2**

* $p < 0.05$; ** $p < 0.005$; *** $p < 0.0005$.

Table 1. Cognitive scores for the dyslexic and control groups (mean and standard deviation). Dyslexics differed from their good-reader peers in all phonological tasks and verbal working memory, but not in their general reasoning skills (Mann-Whitney U tests).

532

533 *Stimuli for dissociating between context effects of recent versus earlier trials -*

534 *Experiment 1*

535 Participants performed 4 blocks of 150 trials of the 2-Tone Frequency Discrimination task.
536 Each trial contained a tone pair (50 ms, 70 dB each tone; 600 ms inter-tone intervals). They were
537 asked to indicate which of the 2 tones had a higher pitch. A short demo of 10 trials preceded the
538 actual experiment. Feedback was provided only in the demo trials. 80% success on the 10 demo
539 (easy) trials was a prerequisite for continuing the task. We did not administer feedback during
540 the assessment because we did not want to affect the magnitude of the listeners' contraction bias.

541 The task was administered with a set of constant stimuli that we designed specifically for this
542 experiment (available at: elsc.huji.ac.il/sites/default/files/glocal_tones.xls). Its design allowed us
543 to evaluate the contribution to the context effect of the most recent trial separately from that of
544 all previous trials. Assessing these effects separately required a specifically designed sequence,
545 since these effects are typically correlated. Specifically, the direction of the frequency distance
546 between the first tone of the current trial and that of the most recent trial, and the direction of the
547 distance from that tone to the average across trials, are typically correlated. In the design of this
548 series we ensured that they were not correlated. In other words, the sign of the global context
549 (G): $G(t) = \text{sign}(f_1(t) - \langle f_1 \rangle)$ and that of the recent context (R):
550 $R(t) = \text{sign}(f_1(t) - f_1(t-1))$, in each trial were not correlated. In this sequence the overall
551 contribution of both the local and global context was positive, yet small.

552

553

Stimuli for assessing the impact of time intervals - Experiments 2a and 2b

In the active condition (Experiment 2a) participants were administered 4 blocks of 100 trials each of the 2-tone frequency discrimination task. Each trial contained a tone pair (50 ms, 70 dB each tone; 600 ms inter-tone intervals), and listeners were asked to indicate which of the 2 tones had a higher pitch. The sequence of trials was randomly drawn for each participant. In each trial, a tone was chosen randomly from 800 Hz to 1,250 Hz. The other tone was chosen randomly-to have a frequency difference (plus or minus) between 1% and 30% from the previously chosen tone. The order of the tones was also randomly chosen. Trial onset asynchrony (TOA) was fixed for each block at 2, 3.5, 6.5 or 9.5 seconds. Block order was counterbalanced across subjects.

In the first, passive, part of the session (Experiment 2b), only the first tone in each pair was presented. We compensated for it by increasing the Inter-Stimulus Intervals (between the tone's offset on the previous trial and the onset of the current trial) in this condition by 0.6 sec. Consequently, the onset-to-onset intervals between first tones of adjacent events were the same in the two conditions. Subjects watched a silent movie and were asked to ignore the tones.

Stimuli for assessing the dynamics of benefits in oral reading rate - Experiment 3

Participants were presented with 6 blocks of 120 non-words and were asked to read them aloud as fast as they could. Voice onset and offset were acquired. Each non-word was presented 500 ms after the voice offset of the preceding non-word. Presentation remained until the voice offset of the current non-word. Non-words were randomly generated by conjunction of two randomly chosen valid Hebrew syllables (consonant-vowel and consonant-vowel-consonant, or vice versa).

ERP recordings and analyses

Electrophysiological activity was recorded from 32 active Ag-AgCl electrodes mounted on an elastic cap using the BioSemi ActiveTwo tools and recording software (BioSemi B.V., Amsterdam, Netherlands). Electrode sites were based on the 10-20 system (American Electroencephalographic Society, 1991). Two additional electrodes were placed over the left and right mastoids. Horizontal EOG was recorded from two electrodes placed at the outer canthi of both eyes. Vertical EOG was recorded from electrodes on the infraorbital and supraorbital regions of the right eye in line with the pupil.

EEG and EOG signals were sampled at 256 Hz, amplified and filtered with an analogue band-pass filter of 0.16 – 100 Hz. Offline analysis was performed using Brain Vision Analyzer 1.05 software (Brain Products GmbH, Gilching, Germany) and EEGLAB toolbox for Matlab (Delorme and Makeig, 2004). The EEG signal was digitally band-pass filtered between 1 Hz and 30 Hz to remove large drifts in signal and high frequency noise. ICA analysis was trained on the entire length of each block and on all scalp electrodes to identify components that reflect eye blink or eye movement evoked electrical activity. An eye related component was identified by its time-correlation with the occurrence of blinks or saccades. This relationship between the identified component and eye blink activity was verified by controlling that the component's scalp distribution was typical of eye-related electrical activity (Delorme et al., 2007). Data were referenced to the nose channel to remove external electrical influence. Artifact rejection was applied to the non-segmented data according to the following criteria: any data point with an EOG or EEG $> \pm 100 \mu\text{V}$ was rejected along with the data ± 300 ms around it. In addition, if the difference between the maximum and the minimum amplitudes of two data points within an interval of 50 ms exceeded $100 \mu\text{V}$, data ± 200 ms around it were rejected. Finally, if the difference between two adjacent data points was more than $50 \mu\text{V}$, the data ± 300 ms around it

were rejected. Trials containing rejected data points were omitted from further analysis. Groups did not differ on the number of trials that were analyzed (active condition: controls: 387 ± 6 ; dyslexics: 382 ± 5 ; $z = 1.6$, *n.s.*; passive condition: controls: 394 ± 4 ; dyslexics: 394 ± 4 ; $z = -0.2$, *n.s.*; mean number of trials \pm SEM, Mann-Whitney U tests)

For ERP averaging across trials, the EEG was parsed to 2,000 ms epochs starting 500 ms before the onset of the first stimulus in each pair, and averaged separately for each electrode. The baseline was adjusted by subtracting the mean amplitude of the pre-stimulus period (500-150 ms before the onset of the first stimulus in the trial) of each ERP from each data point in the epoch. The pre-stimulus baseline period was calculated from this time interval to exclude effects of anticipatory responses that preceded informative anticipated stimuli (CNV; Walter et al., 1964).

ERP analysis was based on the epochs that were recorded with electrode Cz (at the vertex of the scalp). This electrode measured the most prominent response to the auditory stimuli. Data from each acquisition session were analyzed separately. The magnitude of the ERP components was calculated as the area under the curve between 70 ms and 130 ms after first tone's onset for N1 and 150-250 ms time range for P2. We repeated the entire analysis for adjacent electrodes (Fz, FC1, FC2, C3, C4, CP1, CP2 and Pz) and found similar results.

Estimation of the dynamics of recovery from adaptation

Context effects and ERPs in each of the four different ITIs (experiment 2) were fitted with an exponential decay model for each participant separately. This model was previously used to characterize the decay of context effects both behaviorally and for MEG measurements (Lu et al., 1992; Sams et al., 1993). In a previous study, the effect of context as a function of number of trials back was quantitatively measured and indeed resembled an exponential decay (Raviv et al.,

2012). Following that quantitative description, we found that the exponential decay of previous trials captures context effects on this task (Jaffe-Dax et al., 2015). In the current study we modeled the impact of previous trials as a function of the temporal interval instead of number of trials. The model, $\alpha + \beta \exp(-t/\tau)$, had three parameters: α – asymptote after recovery; i.e., the value expected when $t \rightarrow \infty$, β – the magnitude of adaptation; i.e., the value expected at $t = 0$ minus α , τ – time constant of adaptation; i.e., the time it takes for the measure expected at $t = 0$ to decay to $1/e$ (~37%) of its initial value. A small τ indicates fast decay and a large τ indicates a slow decay.

Formally, we searched for the triplet of parameters that minimizes the squared difference between the data and the model prediction: The units of α and β are those of the fitted measure (d' for behavioral bias, $\mu V X msec$ for ERP). In the passive condition, t is the Inter-Stimulus Interval (ISI) between the offset of a tone and the onset of the consecutive tone. In the active condition, t is the Inter-Trial Interval (ITI) between the offset of the second tone and the onset of the first tone in the consecutive trial. Fitted parameters were estimated by minimizing the squared error of the exponential decay model in a limited range. For the ERP adaptation, limits were from 0 to 15,000 for α and from -15,000 to 0 for β . For the d' difference, the limits were from 0 to 100 for α and β . For both measures, τ was limited to be from 0 to 100 seconds. The groups did not differ on the exponential curve's Goodness-of-Fit for any of the analyzed measurements ($z < 1.8$, n.s. Mann-Whitney U tests). The exponential decay model captured the ERP adaptation decay well for both groups ($R^2 > 0.4$). The model was less able to account for the behavioral bias decay, especially for controls ($R^2 = 0.2$), suggesting a more complex mechanism than could be well described by a single exponential decay.

For purposes of compatibility with our previous studies (Raviv et al., 2012; Jaffe-Dax et al., 2015) and to avoid assumptions of normal distribution, we used conservative, non-parametric tests, throughout the study. Using standard parametric tests yielded similar statistical significance.

Conflict of interests

The authors declare no competing financial interests.

Acknowledgements

We thank Ofri Raviv for his useful advice in the creation of the stimuli for Experiment 1 and Elad Goldfarb for his help in verifying word onset times in Experiment 3. This study was supported by the Israel Science Foundation (ISF grant no. 616/11 and Canada-Israel grant no. 2425/15), the Gatsby Charitable Foundation, EPFL-HUJI collaboration, The German-Israeli Foundation for Scientific Research and Development (grant no. I-1303-105.4/2015), Canadian Institutes of Health Research (CIHR), The International Development Research Center (IDRC) and the Azrieli Foundation.

References

- Ahissar M (2007) Dyslexia and the anchoring-deficit hypothesis. *Trends Cogn Sci* 11:458–465
Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17983834> [Accessed January 24, 2011].
- Ahissar M, Lubin Y, Putter-Katz H, Banai K (2006) Dyslexia and the failure to form a perceptual anchor. *Nat Neurosci* 9:1558–1564 Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/17115044> [Accessed March 30, 2012].

Baddeley A (2010) Working memory. *Curr Biol* 20:R136–R140 Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/21961947>.

Banai K, Yifat R (2012) Auditory working memory and early reading skills in Hebrew-speaking preschool children. *J Basic Clin Physiol Pharmacol* 23:109–115 Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/23091274> [Accessed July 13, 2014].

Beneventi H, Tønnessen FE, Ersland L, Hugdahl K (2010) Working memory deficit in dyslexia: behavioral and fMRI evidence. *Int J Neurosci* 120:51–59 Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/20128672> [Accessed October 10, 2010].

Ben-Yehudah G, Ahissar M (2004) Sequential spatial frequency discrimination is consistently impaired among adult dyslexics. *Vision Res* 44:1047–1063 Available at:

[http://www.sciencedirect.com/science/article/B6T0W-4BK2B83-](http://www.sciencedirect.com/science/article/B6T0W-4BK2B83-5/2/b03e863a124766904b5a8048c1d1b4cf)

[5/2/b03e863a124766904b5a8048c1d1b4cf](http://www.sciencedirect.com/science/article/B6T0W-4BK2B83-5/2/b03e863a124766904b5a8048c1d1b4cf).

Bishop DVM, McArthur GM (2004) Immature cortical responses to auditory stimuli in specific language impairment: evidence from ERPs to rapid tone sequences. *Dev Sci* 7:F11–F18

Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15484585>

ISI>://WOS:000223773000001\http://onlinelibrary.wiley.com/store/10.1111/j.1467-

7687.2004.00356.x/asset/j.1467-

7687.2004.00356.x.pdf?v=1&t=hwrwicbp&s=a615b0524267f1b46f17b6a655751e15fd9b1

144.

Boets B, Op de Beeck HP, Vandermosten M, Scott SK, Gillebert CR, Mantini D, Bulthé J,

Sunaert S, Wouters J, Ghesquière P (2013) Intact but less accessible phonetic

representations in adults with dyslexia. *Science* 342:1251–1254 Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/24311693> [Accessed December 11, 2013].

- Boets B, Wouters J, van Wieringen A, Ghesquière P (2007) Auditory processing, speech perception and phonological ability in pre-school children at high-risk for dyslexia: a longitudinal study of the auditory temporal processing theory. *Neuropsychologia* 45:1608–1620 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17303197> [Accessed July 30, 2013].
- Clark KA, Helland T, Specht K, Narr KL, Manis FR, Toga a W, Hugdahl K (2014) Neuroanatomical precursors of dyslexia identified from pre-reading through to age 11. *Brain* 137:3136–3141 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25125610>.
- Daikhin L, Ahissar M (2015) Fast learning of Simple Perceptual Discrimination Reduces Brain Activation in Working Memory and in High-level Auditory Regions. *J Cogn Neurosci* 27:1308–1321.
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134:9–21 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15102499> [Accessed July 4, 2011].
- Delorme A, Sejnowski T, Makeig S (2007) Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage* 34:1443–1449.
- Deutsch A, Bentin S (1996) Attention factors mediating syntactic deficiency in reading-disabled children. *J Exp Child Psychol* 63:386–415.
- Díaz B, Hintz F, Kiebel SJ, von Kriegstein K (2012) Dysfunction of the auditory thalamus in developmental dyslexia. *Proc Natl Acad Sci U S A* 109:13841–13846 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3427071&tool=pmcentrez&rendertype=abstract> [Accessed August 13, 2013].
- Fischer J, Whitney D (2014) Serial dependence in visual perception. *Nat Neurosci* 17:738–747 Available at: <http://dx.doi.org/10.1038/nn.3689>.

Hämäläinen J a, Salminen HK, Leppänen PHT (2013) Basic Auditory Processing Deficits in Dyslexia: Systematic Review of the Behavioral and Event-Related Potential/Field Evidence. *J Learn Disabil* 46:413–427 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22323280> [Accessed August 4, 2013].

Hari R, Kaila K, Katila T, Tuomisto T, Varpula T (1982) Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: implication for their neural generation. *Electroencephalogr Clin Neurophysiol* 54:561–569.

Herrmann B, Henry MJ, Fromboluti EK, McAuley JD, Obleser J (2015) Statistical context shapes stimulus-specific adaptation in human auditory cortex. *J Neurophysiol* 113:2582–2591 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25652920> <http://dx.doi.org/10.1152/jn.00634.2014>.

Hornickel J, Kraus N (2013) Unstable Representation of Sound: A Biological Marker of Dyslexia. *J Neurosci* 33:3500–3504 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4205-12.2013> [Accessed February 21, 2013].

Jaffe-Dax S, Raviv O, Jacoby N, Loewenstein Y, Ahissar M (2015) A Computational Model of Implicit Memory Captures Dyslexics' Perceptual Deficits. *J Neurosci* 35:12116–12126 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1302-15.2015>.

Kleiner M, Brainard D, Pelli D (2007) What's new in Psychtoolbox-3? In: *Perception* 36 ECVP.

Landerl K, Wimmer H, Frith U (1997) The impact of orthographic consistency on dyslexia : A German-English comparison. *Cognition* 63:315–334.

Lanting CP, Briley PM, Sumner CJ, Krumbholz K (2013) Mechanisms of adaptation in human auditory cortex. *J Neurophysiol* 110:973–983 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3742970&tool=pmcentrez&ren>

dertype=abstract.

Lu Z-L, Williamson J, Kaufman L (1992) Behavioral Lifetime of Human Auditory Sensory

Memory Predicted by Physiological Measures. *Science* (80-) 258:1668–1670.

MacKay DG (1970) Spoonerisms: the structure of errors in the serial order of speech.

Neuropsychologia 8:323–350 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/5522566>.

Macmillan NA, Creelman CD (2004) *Detection Theory: A User's Guide*, 2nd ed. Taylor &

Francis.

Mayhew SD, Dirckx SG, Niazy RK, Iannetti GD, Wise RG (2010) EEG signatures of auditory

activity correlate with simultaneously recorded fMRI responses in humans. *Neuroimage*

49:849–864 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19591945> [Accessed

September 12, 2011].

McAnally KI, Stein JF (1996) Auditory Temporal Coding in Dyslexia. *Proc R Soc B Biol Sci*

263:961–965 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8805833> [Accessed

November 6, 2011].

McArthur GM, Bishop DVM (2005) Speech and non-speech processing in people with specific

language impairment: a behavioural and electrophysiological study. *Brain Lang* 94:260–273

Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16098376>.

Micheyl C, Delhommeau K, Perrot X, Oxenham AJ (2006) Influence of musical and

psychoacoustical training on pitch discrimination. *Hear Res* 219:36–47 Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/16839723>.

Möller J, Jansma BM, Rodriguez-Fornells A, Münte TF (2007) What the brain does before the

tongue slips. *Cereb cortex* 17:1173–1178 Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/16831855> [Accessed September 17, 2013].

Nicolson RI, Fawcett AJ, Brookes RL, Needle J (2010) Procedural Learning and Dyslexia.

Dyslexia 16:194–212.

Oganian Y, Ahissar M (2012) Poor anchoring limits dyslexics' perceptual, memory, and reading skills. *Neuropsychologia* 50:1895–1905 Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0028393212001650> [Accessed May 7, 2012].

Parbery-Clark A, Strait DL, Kraus N (2011) Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia* 49:3338–3345 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3445334&tool=pmcentrez&rendertype=abstract> [Accessed August 13, 2013].

Preuschhof C, Schubert T, Villringer A, Heekeren HR (2010) Prior Information biases stimulus representations during vibrotactile decision making. *J Cogn Neurosci* 22:875–887 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19413475>.

Pugh KR, Frost SJ, Sandak R, Landi N, Rueckl JG, Constable RT, Seidenberg MS, Fulbright RK, Katz L, Mencl WE (2008) Effects of stimulus difficulty and repetition on printed word identification: an fMRI comparison of nonimpaired and reading-disabled adolescent cohorts. *J Cogn Neurosci* 20:1146–1160 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18284344>.

Ramus F (2014) Neuroimaging sheds new light on the phonological deficit in dyslexia. *Trends Cogn Sci*:3–6 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24568928>.

Ramus F, Ahissar M (2012) Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance. *Cogn Neuropsychol* 29:37–41 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22559749> [Accessed July 18, 2012].

Raviv O, Ahissar M, Loewenstein Y (2012) How recent history affects perception: the normative approach and its heuristic approximation. *PLoS Comput Biol* 8:e1002731 Available at:

<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3486920&tool=pmcentrez&rendertype=abstract> [Accessed February 11, 2013].

Raviv O, Lieder I, Loewenstein Y, Ahissar M (2014) Contradictory Behavioral Biases Result from the Influence of Past Stimuli on Perception. *PLoS Comput Biol* 10:e1003948 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25474117> [Accessed December 5, 2014].

Sams M, Hari R, Rif J, Knuutila J (1993) The Human Auditory Sensory Memory Trace Persists about 10 sec: Neuromagnetic Evidence. *J Cogn Neurosci* 5:363–370.

Schab FR, Crowder RG (1988) The role of succession in temporal cognition: is the time-order error a recency effect of memory? *Percept Psychophys* 44:233–242 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3174355>.

Sheehan KA, McArthur GM, Bishop DVM (2005) Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Brain Res Cogn Brain Res* 25:547–553 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16198089> [Accessed November 12, 2010].

Stewart N, Brown GD a (2004) Sequence effects in the categorization of tones varying in frequency. *J Exp Psychol Learn Mem Cogn* 30:416–430.

Tallal P (1980) Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang* 9:182–198 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7363063>.

Temple E, Poldrack R a, Protopapas a, Nagarajan S, Salz T, Tallal P, Merzenich MM, Gabrieli JD (2000) Disruption of the neural response to rapid acoustic stimuli in dyslexia: evidence from functional MRI. *Proc Natl Acad Sci U S A* 97:13907–13912 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=17674&tool=pmcentrez&rendertype=abstract>.

- Torgeson J, Goldman T (1977) Verbal rehearsal and short-term memory in reading-disabled children. *Child Dev* 48:55–60 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/844361>.
- Tremblay KL, Inoue K, McClannahan K, Ross B (2010) Repeated stimulus exposure alters the way sound is encoded in the human brain. *PLoS One* 5:e10283 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2858650&tool=pmcentrez&rendertype=abstract> [Accessed February 5, 2014].
- Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12652303> [Accessed July 24, 2010].
- Walter WG, Cooper R, Aldridge VJ, Mccallum WC, Winter a L (1964) Contingent Negative Variation: an Electric Sign of Sensorimotor Association and Expectancy in the Human Brain. *Nature* 203:380–384.
- Wechsler D (1997) Wechsler Adult Intelligence Scale (WAIS-III) Administration and Scoring Manual. San Antonio, TX: The Psychological Corporation.
- Wei X-X, Stocker AA (2015) A Bayesian observer model constrained by efficient coding can explain “anti-Bayesian” percepts. *Nat Neurosci* 18:1509–1517 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26343249>.
- Witton C, Talcott JB, Hansen PC, Richardson a J, Griffiths TD, Rees a, Stein JF, Green GG (1998) Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Curr Biol* 8:791–797 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9663387>.
- Woodrow H (1933) Weight-Discrimination with a Varying Standard. *Am J Psychol* 45:391–416.
- Zoccolotti P, Friedmann N (2010) From dyslexia to dyslexias, from dysgraphia to dysgraphias, from a cause to causes: A look at current research on developmental dyslexia and

835 dysgraphia. Cortex 46:1211–1215 Available at:
836 <http://www.ncbi.nlm.nih.gov/pubmed/20933228> [Accessed October 27, 2010].
837