Supplementary materials

1. Behavioural tasks

1.1. Reading behaviour

Word and pseudo-word reading times and error rates were recorded. The stimuli were highly concrete and familiar nouns. There were 20 two-syllable words and 20 pseudo-words. Pseudo-words were created from the 20 words by maintaining the 'word envelope' while changing internal consonants. Each stimulus was displayed for a maximum of 1.5 seconds. Participants were asked to read each word or pseudo-word as soon as it appeared on the screen. Voice onset times for detection of a single dot were used as a baseline measure to further evaluate the performance in the reading task: this reference measure allowed us to distinguish a specific reading disorder from a generalized lengthening of vocal reaction times for visual stimuli (Paulesu *et al.*, 2001).

The tasks were adminstered by a HP (Pavillon zv50000) computer using the software SuperLab.

1.2. Phonological tasks

<u>Spoonerism task</u>

Subjects heard pairs of words with the instruction to repeat back the two words after having swapped the initial sound around (e.g., **C**ASA and **T**OPO repeated as **T**ASA and **C**OPO). The 24 words used in the Spoonerism task were highly familiar concrete words with two syllables. They were selected on the basis that they had clear syllable divisions and no consonant clusters in their onset. Performance was measured as the cumulative time taken to produce the 12 spoonerisms. Time recording for each stimulus started after its presentation. Error rate was also recorded (Perin, 1983).

<u>Digit naming task</u>

In this task, participants were asked to read aloud as fast as possible 100 digits, divided in two blocks of 50. Within each block, digits were chunked into strings of 5 (e.g., 68248 83542 99634); participants were asked to read each digit of a given string individually: for example, the string 51368 was read as «cinque, uno, tre, sei, otto». The time taken to read each block of 50 digits was recorded and a mean score, in seconds, was obtained over the two blocks (Denckla and Rudel, 1976).

<u>Picture naming task</u>

Subjects named 7 pictures as fast as possible. To minimize the strategic aspects of word searching or any variability due to the word used to name a given picture (e.g., a computer can be named "computer" as well as "PC"), subjects were familiarized with the task several times (Paulesu *et al.*, 2001). VOTs and error rates were recorded. This task was also administered using a HP (Pavillon zv50000) computer with the software SuperLab.

1.3. "Dorsal visual stream skills" *tasks*

Three different tasks were used following the procedures described in Ramus *et al.* (2003).

Speed Discrimination

The stimuli used in this experiment were Gabor patterns: cosinusoidal vertical gratings spatially windowed by an isotropic Gaussian. More in details the stimuli are represented as a function s(x,y,t) which depends on the spatial coordinates (x,y) and on the time, t, as follows:

$$s(x,y,t) = \frac{\Delta L}{2}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[\frac{x^2+y^2}{2\sigma^2}\right]\cos\left[\frac{2\pi}{\lambda_s}x - \omega t + \varphi\right] + L_{min} \qquad (1)$$

where L_{min} is the minimum luminance and ΔL represents the maximum variation. Their value is computed from the mean luminance of the stimulus, which is equal to 5 cd/m^2 with range from 0 cd/m^2 to 10 cd/m^2 for the magno (M-) selective stimuli, and to 20 cd/m^2 from 0 cd/m^2 to 40 cd/m^2 for the parvo (P-) selective stimuli. In the actual stimuli, the range was restricted to 80% of the full range: from 1 cd/m^2 to 9 cd/m^2 for the magno (M-) stimuli and from 4 cd/m^2 to 36 cd/m^2 for the parvo (P-) stimuli

The reference horizontal angular drifting velocity ω. was set to $\overline{\omega} = 1°/s$ for the M-selective case and $\overline{\omega} = 16°/s$ for the P-selective case. The value of the wavelength λ_s was set equal to 2° and to 16° for the M- and P- selective stimuli respectively, and the phase φ was set randomly between 0 and 2π in each trial.

Finally in order to prevent subjects counting the number of bars passing rather than judging speed, the signal s(.) is ramped on and off with a Gaussian contrast time envelop with a standard deviation randomly chosen in the range between 160 and 240 msec, with uniform distribution. The duration of the entire stimulus was set to 500 msec.

Subjects were presented a certain number of trials each of them with two stimuli: one with the reference speed and the other with a larger speed. The order of the two stimuli is randomly selected. The subjects' task was to indicate which stimulus was faster.

Threshold adaptation

QUEST (Watson and Pelli, 1983) was used to estimate the percentage increase in speed of the stimulus with respect to the fixed baseline speed ($\overline{\omega} = 1°/s$ for the M-selective case and $\overline{\omega} = 16°/s$ for the P-selective case), required to discriminate the two stimuli with an expected 83% of accuracy. To this aim we continuously monitor the response of the subject and correct the speed increment of each next trial based on their actual response and to all previous responses.

More in details, we started with the hypothesis that all trials are independent. We obtain this by interleaving the trials targeted on Magno and Parvo cellular systems.

We then represented the information in the response sequence obtained up to a given trial, in the form of a likelihood function, which represents the probability to have the sequence of response measured, A^m, given a certain threshold T:

$$f_{A^m|T}(A^m|T) = \prod_{i=1}^{n} \Pr(A_i^m(x_i)|T) \tag{2}$$

where n is the number of trials and x_i is equal to the speed increment, $\Delta\omega$. Since each trial is a 2 Alternative Forced Choice task, the single answer, A_i^m, associated to a given stimulus intensity x_i, is formally represented by a Bernoulli random variable whose possible values are {*correct*, *not correct*} and whose probability is given by:

$$\Pr(A_i^m(x_i)|T) = \begin{cases} \psi(x_i;T) & A_i^m \text{ is correct} \\ 1-\psi(\rho_i;T) & \text{otherwise} \end{cases} \tag{3}$$

where the function $\psi(\cdot\,;T)$ is the psychometric function. This has a typical sigmoidal shape, and it can be expressed by a Weibull function, given by:

$$\psi(x;T) = 1-(1-\gamma)\exp\left[-10^{\left(\frac{\beta}{20}\right)(x-T+\varepsilon)}\right] \tag{4}$$

where the intensity and the threshold are expressed in decibel (we will indicate x as the log-intensity and T as the log-threshold).

γ is a small value that corresponds to the probability of failing when the task is very easy. It could be due to a mental lapse, for example, β is the slope of the performance function; and ε is the criterion that is adjusted so that $\psi(0;T) = p_{des}$. In our case, γ was set to a very small value of .001. β

was set to *3.5* as suggested in [P. King-Smith, S. Grigsby, A. Vingrys, S. Benes, and A. Supowit. "Efficient and unbiased modifications of the quest threshold method: Theory, simulations, experimental evaluation and practical implementation." Vision Res., vol. 34, no. 7, pp. 885 – 912. 1994] and $\varepsilon = .1085$, which corresponds to a desired success rate p_{des} of 80%.
Changing *T*, the curve is shifted horizontally.

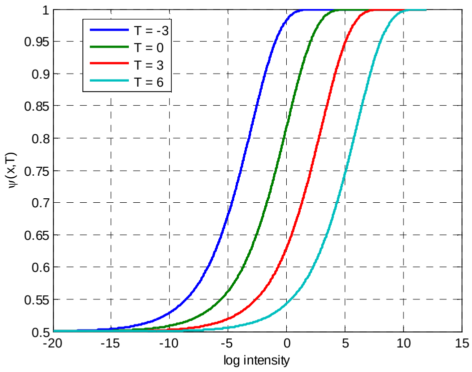

Figure 1

The Weibull distribution for different threshold values

We can also consider the results of experiments reported in the literature that have suggested an adequate threshold value for this experiment. We take this value inside a statistical framework and express it as a Gibbs potential with a Gaussian form:

$$f_T(T) = \frac{1}{\sqrt{2\pi}\sigma_T} \exp\left[-\frac{1}{2}\left(\frac{T-\mu_T}{\sigma_T}\right)^2\right] \tag{5}$$

where $\mu_T = 50$ for the Magno case and $\mu_T = 80$ for the Parvo case, according to (Ramus *et al.*, 2003).

These two pieces of information can be grouped together into a Maximum A Posteriori (MAP) estimate that takes the expression of a conditional probability of the threshold given the data and the a-priori information:

$$f_{T|A^m}(T|A^m) = \frac{f_{A^m|T}(A^m|T)\, f_T(T)}{f_{A^m}(A^m)} \tag{6}$$

The function $f_{A^m}(A^m)$ is the prior density probability of the data. It does not depend on the threshold and it plays the role of a normalization factor and is usually discarded.

The posterior probability density function (6) contains all the available information about the threshold: assumptions, prior estimate and data, and therefore it can be used to set the value of the next threshold in each trial. Maximizing Eq. (6) the best estimate of the next threshold is obtained.

To achieve this, we derive from Eq. (6) the *quest function* transforming it into logarithms:

$$Q(T) = \ln f_{T,A^m}(T, A^m) = \ln f_{A^m|T}(A^m|T) + \ln f_T(T) \tag{7}$$

We introduce the Bernoulli likelihood (Eq. (3)) into (7) to obtain:

$$Q(T) = \ln f_{T,A}(T, A^m) = \sum_{i=1}^{n} \ln \Pr(A_i^m(\rho_i)|T) + \ln f_T(T) \tag{8}$$

Here we suppose that speed variations are inside a given range (here from 28°/s to 89°/s for the M-stimuli and from 45°/s to 142°/s for the P-stimuli) and they are spaced at equal intervals, Δx (the grain) along the log intensity axis. Therefore, $Q(T)$ could be evaluated only at $x_j = x_0 + j\ \Delta x$ values, with j integer.

So, at each iteration, n, a vector containing the values of Q for all the possible x_j, is computed as follows:

$$Q_n(x_j) = \ln f_T(x_j) + \sum_{i=1}^{n} \ln \Pr(A_i(\rho_i)|x_j). \tag{9}$$

The MAP estimate at the trial n is compute as the maximum of the vector Q_n, and this value will be used as threshold for the trial n+1.

An efficient way of computing Eq. (9) can be derived rewriting it in recursive way as:

$$Q_n(x_j) = \ln f_T(x_j) + \sum_{i=1}^{n-1} \ln \Pr(A_i^m(\rho_i)|x_j) + \ln \Pr(A_n^m(\rho_n)|T) = Q_{n-1}(x_j) + \ln \Pr(A_n^m(\rho_n)|x_j). \tag{10}$$

If we indicate:

$$\psi(x) = \psi(x; 0) \tag{11}$$

it follows that:

$$\psi(x - T) = \psi(x; T). \tag{12}$$

So, when the n^{th} trial is performed, with a chosen threshold $\hat{x}$ (it will be one of the x_js), the update of $Q_n(\cdot)$ becomes:

$$Q_n(x_j) = Q_{n-1}(x_j) + \begin{cases} \ln \psi(x_j - \hat{x}) & A_n^m \text{ is correct} \\ \ln[1 - \psi(x_j - \hat{x})] & \text{otherwise} \end{cases}. \tag{13}$$

which suggests to pre-compute the two functions $\psi(\cdot)$ and $1 - \psi(\cdot)$ for all possible values, x_j for a faster computation.

The final estimate of the threshold is computed as the maximum likelihood estimate after the last trial as maximum a-posteriori estimate cannot be applied in this particular case.

Contrast Discrimination

The stimuli used in this experiment were Gabor patterns.

Stimuli were cosinusoidal vertical gratings spatially windowed by an isotropic Gaussian. As in Ramus *et al.* (2003), we tested the contrast sensitivity for magnocellular-selective stimuli and for parvocellular-selective stimuli. For the magnocellular case the peak spatial frequency was set to .5 c/deg and counter-phase flickered to a rate of 15 reversals per second; while for the parvocellular case the peak spatial frequency was set to 8.0 c/deg without counter-phase flicker.

In order to minimize the impact of onset and offset transients, the parvocellular stimuli were smoothly ramped on and off with a Gaussian contrast envelope (200 msec), while the contrast of magnocellular stimuli was modulated using a Gabor temporal envelope multiplied by a sinusoid.

Subjects were presented trials containing two stimuli: one contained a Gabor pattern and the other a blank field at background luminance. The order of the two stimuli was randomly selected. Stimulus duration was 500 msec. The subjects' task was to indicate which stimulus contained the Gabor pattern. An auditory cue was used to indicate the onset of each trial. The task was performed by randomizing magnocellular-selective trials and parvocellular-selective trials.

The Bayesian adaptive psychometric method QUEST (Watson and Pelli, 1983) was used to estimate the percentage increase in luminance of the stimulus with respect to the fixed baseline required to discriminate the two stimuli from the background with an expected 83% of accuracy. To this end, subjects' responses were continuously monitored and luminance was modified for each trial on the basis of subjects' actual answers and all previous answers. The number of trials varied depending on how fast subjects were in reaching the completion criteria imposed by the QUEST procedure.

Coherent Motion Perception

Subjects were presented with an 8°×8° field of 150 white dots with randomized position presented on a grey (50 cd/m2) background, Some of these dots moved rapidly (11 °/sec) to the right or to the left, while the remaining dots moved in a random direction.

The subjects' task was to indicate, on average, the direction of motion (left or right). Stimulus duration was 900 msec.

QUEST (Watson and Pelli, 1983) was used to estimate the percentage increase in the proportion of dots moving in a random direction. This involved estimating the minimum proportion of dots moving in coherent direction that was required to discriminate the average direction of stimuli with 83% accuracy. To this aim we continuously monitored the subject's response and adjusted the percentage of dots moving coherently on the next trial, based on all previous responses.

The tasks were administered by a MacMini computer using the software Matlab R2010a (Mathworks®). Stimuli were displayed on a 17" SyncMaster 720N Samsung Monitor operating at screen resolution 800x600 pixels with a frame refresh rate of 85 Hz. Subjects viewed the screen binocularly at a distance of 80 cm in a dark room.

1.4. Cerebellar task

Motor sequence learning

Participants were instructed to identify a sequence of eight taps through a trial-and-error approach using a four key response-box during the fMRI session (Nicolson *et al.*, 1999). In particular, they received a high-pitch acoustic feedback for correct responses and a low pitch for wrong responses. Moreover, they had to return to the beginning of the sequence after each wrong tap. Correct taps and lack of corrections (perseveration on errors) were used as indices of learning.

2. Description of fMRI protocol

The rationale for the selection of these tasks was first and foremost to include "reading" which is the defining test for dyslexia. Because bi-syllabic noun reading is a relatively non-demanding task in a transparent orthography like the Italian one, we opted for pseudo-word reading of bi-syllabic stimuli to better challenge the reading system. The other tasks were selected to challenge, in the most parsimonious way, the other three systems called into play by current theories of dyslexia: phonological processing was assessed with auditory presentation of the stimuli, the visual magnocellular system was assessed with visual motion stimulation, and cerebellar function was assessed with motor sequence learning tasks.

2.1. Pseudo-word reading

The subjects were presented with alternating 30 second blocks of bi-syllabic and tri-syllabic high-frequency pseudo-words (experimental condition). or strings of differently oriented segments covering the same visual angle of the pseudo-words (baseline condition). Stimuli were presented in the centre of the screen at a rate of one every two seconds. The pseudo-words were derived by high-frequency Italian words. The subjects were instructed to covertly read the pseudo-words and to fixate the strings of segments, pressing a button with the right index after the processing of each stimulus (Paulesu *et al.*, 2001).

2.2. Auditory letter-name rhyming task

This simple phonological awareness task implied the assessment of whether a syllable (e.g., [pi], [di], [èf-fe], [ʒè-ta]) rhymed with the syllable [bi] (the same sound of the letter-name B). The baseline condition implied a (same/different) discrimination task with pairs of pure tones. Subjects were presented with alternating 30 seconds blocks of tones and syllables and they were instructed to press a key on a response box using the right index finger when syllables rhymed and when pure tones matched. Presentation rate was 1 stimulus every two and a half seconds, such that every block had 12 stimuli. Target rate was 5/12.

2.3. Visual motion perception task

A low spatial frequency Gabor patch was presented on the computer screen as either stationary (control epochs) or moving randomly across the screen in 4 directions (from the left to the right of the screen horizontally, from the top to the bottom vertically, from the lower left to the upper right in diagonal and vice versa on the other diagonal) for a total of 30 times in each block. Thus, there was a direction-shift for each 500 msec. Goggles simulated a distance of 60 cm from the screen. During the baseline and experimental epochs, the subjects were instructed to focus on a fixation cross at the centre of the screen during the whole scan session, and not to not follow the moving stimuli during the experimental epochs.

2.4. Motor sequence "cerebellar" learning task

The subjects were instructed to identify and to learn a sequence of 8 key-presses using a four key response-box; they received an acute auditory feedback for the correct taps and a low auditory feedback for the wrong taps. In the baseline, they were instructed to passively listen to the same sounds used as feedback during the motor learning task. The subjects were presented with alternating 30-second blocks of baseline and the motor sequence learning task. The rate of the task was set so that the subject had to press a key every 2 seconds.

References

Denckla MB, Rudel RG. Rapid "automatized" naming (R.A.N): dyslexia differentiated from other learning disabilities. Neuropsychologia 1976; 14(4): 471-9.

Nicolson RI, Fawcett AJ, Berry EL, Jenkins IH, Dean P, Brooks DJ. Association of abnormal cerebellar activation with motor learning difficulties in dyslexic adults. Lancet 1999; 353(9165): 1662-7.

Paulesu E, Démonet JF, Fazio F, McCrory E, Chanoine V, Brunswick N, *et al.* Dyslexia: cultural diversity and biological unity. Science 2001; 291(5511): 2165-7.

Perin D. Phonemic Segmentation and Spelling. British Journal of Psychology 1983; 74: 129-44.

Ramus F, Rosen S, Dakin SC, Day BL, Castellote JM, White S, *et al.* Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. Brain 2003; 126(Pt 4): 841-65.

Watson AB, Pelli DG. QUEST: a Bayesian adaptive psychometric method. Percept Psychophys 1983; 33(2): 113-20.

Table sm-1. Spearmann's rho correlation matrices for the phonological, lexical, and motor variables. Bartlet's sphericity test: $Chi^2_{(66)} = 111.95$; $p < .001$. Red highlighted correlations: $p < .05$

	Verbal I.Q.	Performance I.Q.	Voice-Onset Time	Word reading (time)	Pseudo-word reading (time)	Picture naming (time)	Syllable rhyming (d-prime)	Spoonerism (time)	Digit naming (time)	Motor learning (correct taps)	Motor learning (lack of correction)
Verbal I.Q.	1.00	.5	-.11	-.33	-.31	-.16	.27	-.26	-.34	.24	-.39
Performance I.Q.	.5	1.00	.21	.18	-.01	-.06	-.10	-.06	-.10	.16	-.27
Voice-Onset Time	-.11	.21	1.00	.71	.17	.37	-.35	.01	.01	-.34	.06
Word reading (time)	-.33	.18	.71	1.00	.63	.51	-.63	-.02	-.02	-.28	-.02
Pseudo-word reading (time)	-.31	-.01	.17	.63	1.00	.42	-.31	.08	.19	.05	-.16
Picture naming (time)	-.16	-.06	.37	.51	.42	1.00	-.38	.30	.15	.01	-.03
Syllable rhyming (d-prime)	.27	-.10	-.35	-.63	-.31	-.38	1.00	.22	.27	.25	.07
Spoonerism (time)	-.26	-.06	.01	-.02	.08	.30	.22	1.00	.78	.20	.06
Digit naming (time)	-.34	-.10	.01	-.02	.19	.15	.27	.78	1.00	.40	-.09
Motor learning (correct taps)	.24	.16	-.34	-.28	.05	.01	.25	.20	.40	1.00	-.66
Motor learning (lack of correction)	-.39	-.27	.06	-.02	-.16	-.03	.07	.06	-.09	-.66	1.00

Table sm-2. Spearmann's rho correlation matrices for visuo-magnocellular tasks. Bartlet's sphericity test: $Chi^2_{(10)} = 20.44$, $p < .02$; Red highlighted correlations: $p < .05$

	Magnocellular CS (% of increased contrast)	Parvocellular CS (% of increased contrast)	Magnocellular SD (% of increased speed)	Parvocellular SD (% of increased speed)	Coherent motion perception (% of increased proportion of moving dots)
Magnocellular CS (% of increased contrast)	1.00	-.49	.42	.03	.48
Parvocellular CS (% of increased contrast)	-.49	1.00	-.11	-.23	-.28
Magnocellular SD (% of increased speed)	.42	-.11	1.00	-.10	.47
Parvocellular SD (% of increased speed)	.03	-.23	-.10	1.00	-.23
Coherent motion perception (% of increased proportion of moving dots)	.48	-.28	.47	-.23	1.00