Supplementary Information

Table S1. Description of standardized tests

Variable name	Test battery	Task	Example
Word Identification	WJ-III	Read aloud words of varying frequency, regularity and length (unlimited time)	M … for … night … could … spoil …
Word Attack	WJ-III	Read aloud pronounceable pseudowords (unlimited time)	dap… vek… koarb… knofe…
Basic Reading	WJ-III	Composite score: [Word Identification, Word Attack]	
Sight Word Efficiency	TOWRE	Read aloud as many words as you can in a minute	
Phonological Decoding Efficiency	TOWRE	Read aloud as many pseudowords as you can in a minute	
Elision	CTOPP	Remove a phoneme from an aurally presented word, say the remaining word.	Say 'brand'. Now say brand without saying /r/. (band)
Blending	CTOPP	Combine sounds to form a word	What word do these sounds make? /b/-/oi/'. (boy)
Rapid Digit Naming	CTOPP	Name an array of 36 visual digits as fast as possible (Total time for two arrays)	
Rapid Letter Naming	CTOPP	Name an array of 36 visual letters as fast as possible (Total time for two arrays)	

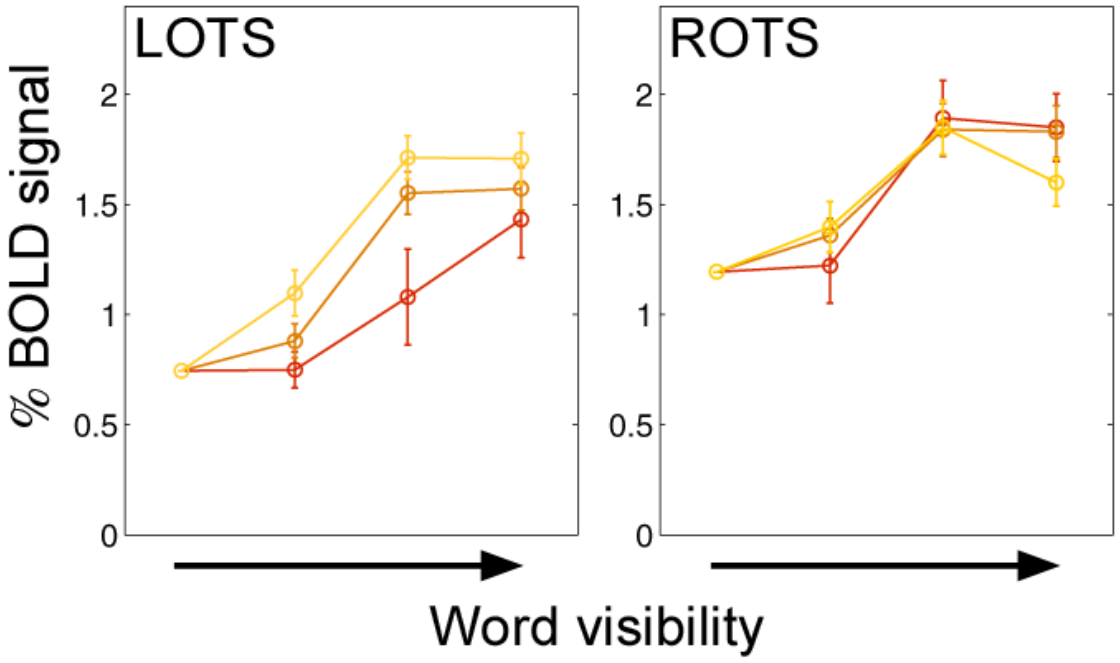

Figure S1. Age grouped neurometric functions in LOTS (left panel) and ROTS (right panel), without smooth curve fitting. BOLD signal is scaled to the mean response for the full noise condition, to reflect overall differences in signal between the regions. Data are grouped by participants' age: 7-8 years old (Brown, N=7), 9-11 years old (orange, LOTS: N=32; ROTS: N=30) and 12-15 years old (yellow, LOTS: N=27; ROTS: N=25). Vertical error bars represent the standard error of the mean computed at each visibility level across the data of participants within that age group. Some participants contributed more than one curve for each age group, for example, a participant who's entry age was 7 contributed two curves to the youngest age group. However, this participant was counted only once for the purpose of calculating the standard error of the mean depicted in the error bars (the Ns above represent the number of different subjects within age group). If anything, the error bars are an over estimation of the real variance.

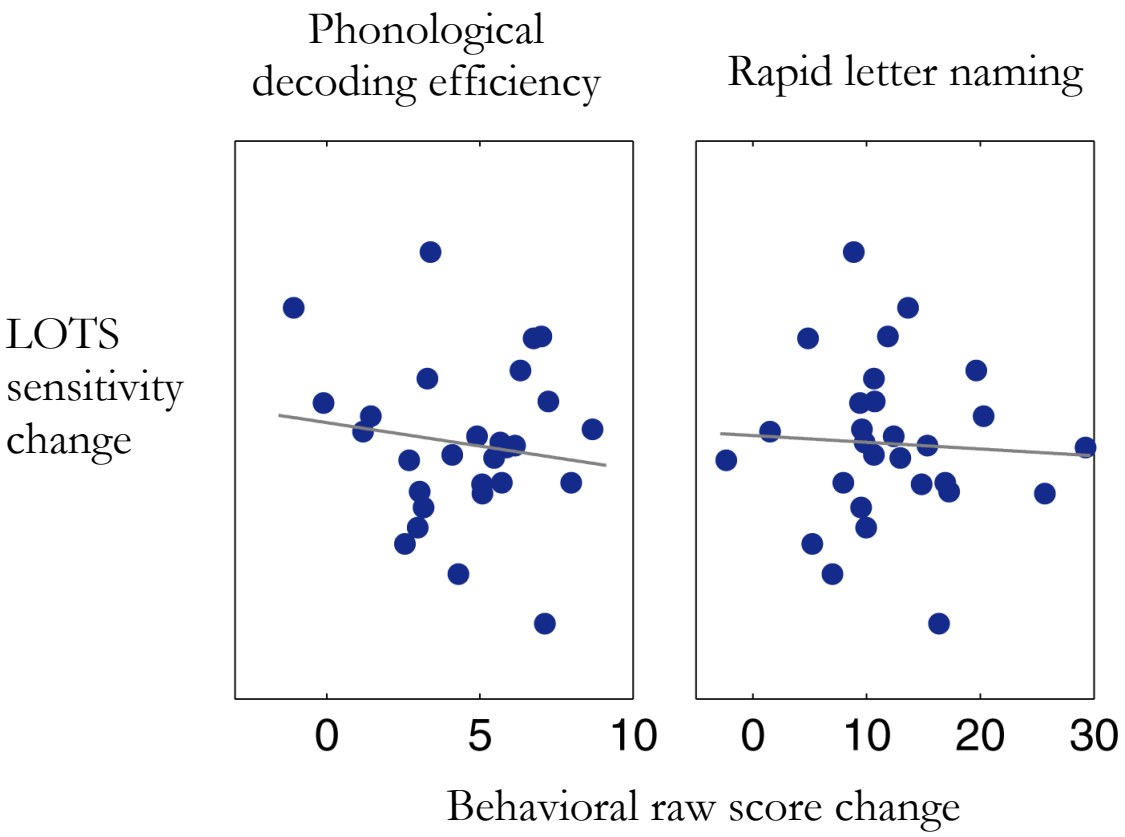

Figure S2. Sensitivity change in LOTS does not correlate with phonological or rapid naming measures. Refer to Figure 2 of the main paper for a graphical explanation of the sensitivity change computation. The two scatter plots parallel Figure 2C, using the phonological decoding efficiency score (TOWRE) and rapid letter naming score (CTOPP). N=28. Rapid letter naming raw scores (the time it takes to name a fixed number of letters) were transformed to reflect the number of letters named per minute, for consistency with the TOWRE measures.

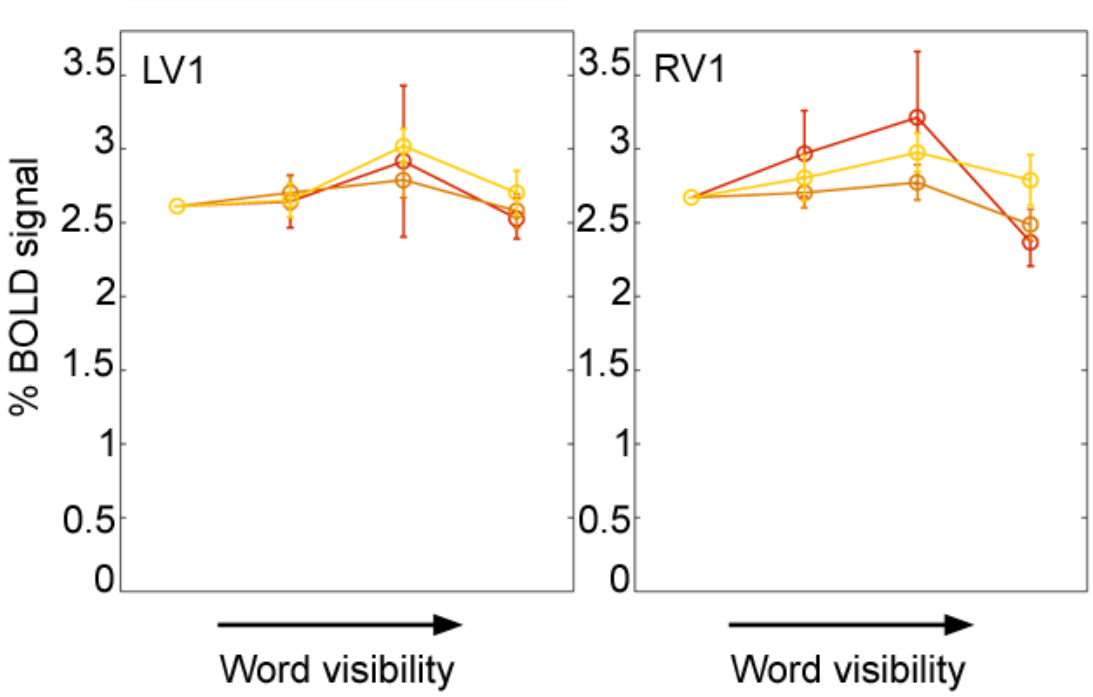

Figure S3. Age-grouped neurometric functions in foveal V1. Data from left V1 (left panel) and right V1 (right panel) were scaled and age-grouped as in Figure S1. While bilateral foveal V1 responded strongly to stimuli at all visibility levels, it showed no developmental trend and no sensitivity to the word visibility manipulation. LV1: N=7, 29, 25; RV1: N= 7, 32, 27.

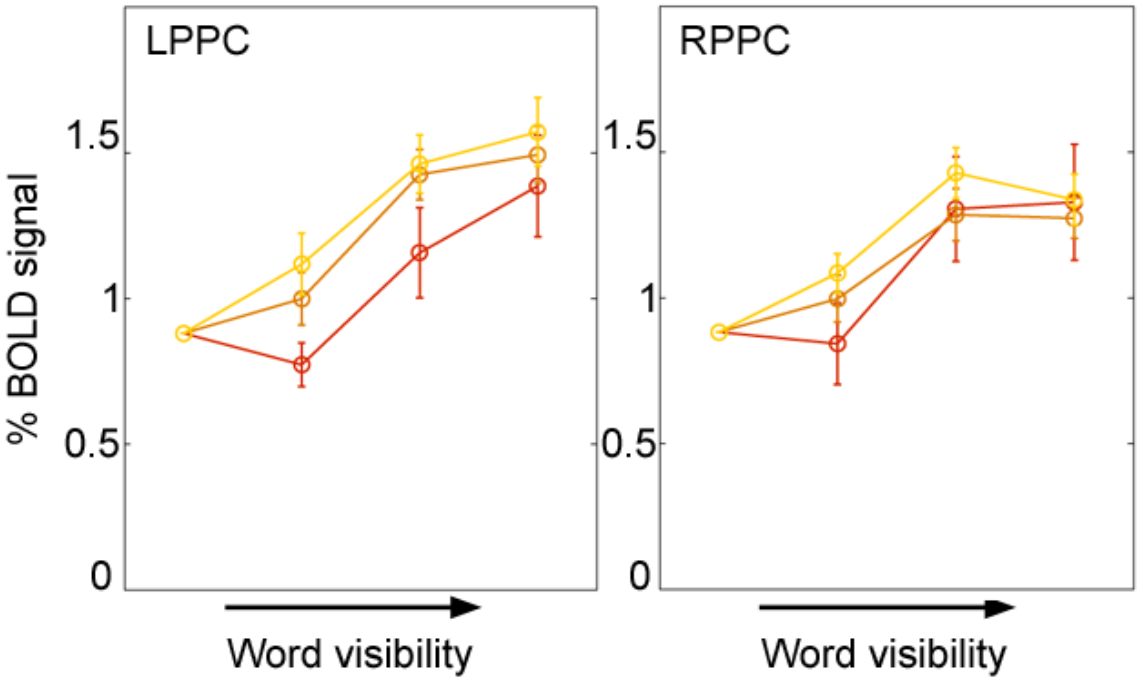

Figure S4. Age-grouped neurometric functions in posterior parietal cortex. Data from left posterior parietal cortex (LPPC, left panel) and right PPC (right panel) were scaled and grouped as in Figure S1, S3. Both regions showed rising neurometric curves, and some developmental trend is evident in the left PPC. N=6, 30, 25. See main text for discussion.

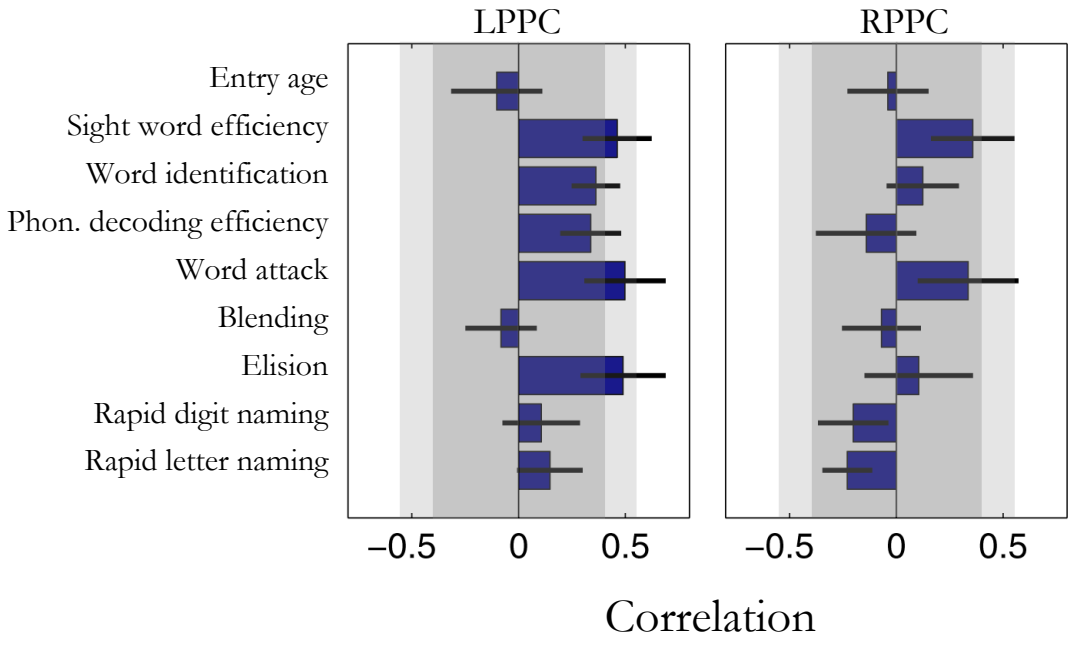

Figure S5. Correlation values between behavioral change and cortical sensitivity change in left and right posterior parietal cortex. See Caption to Figure 3. N=25.

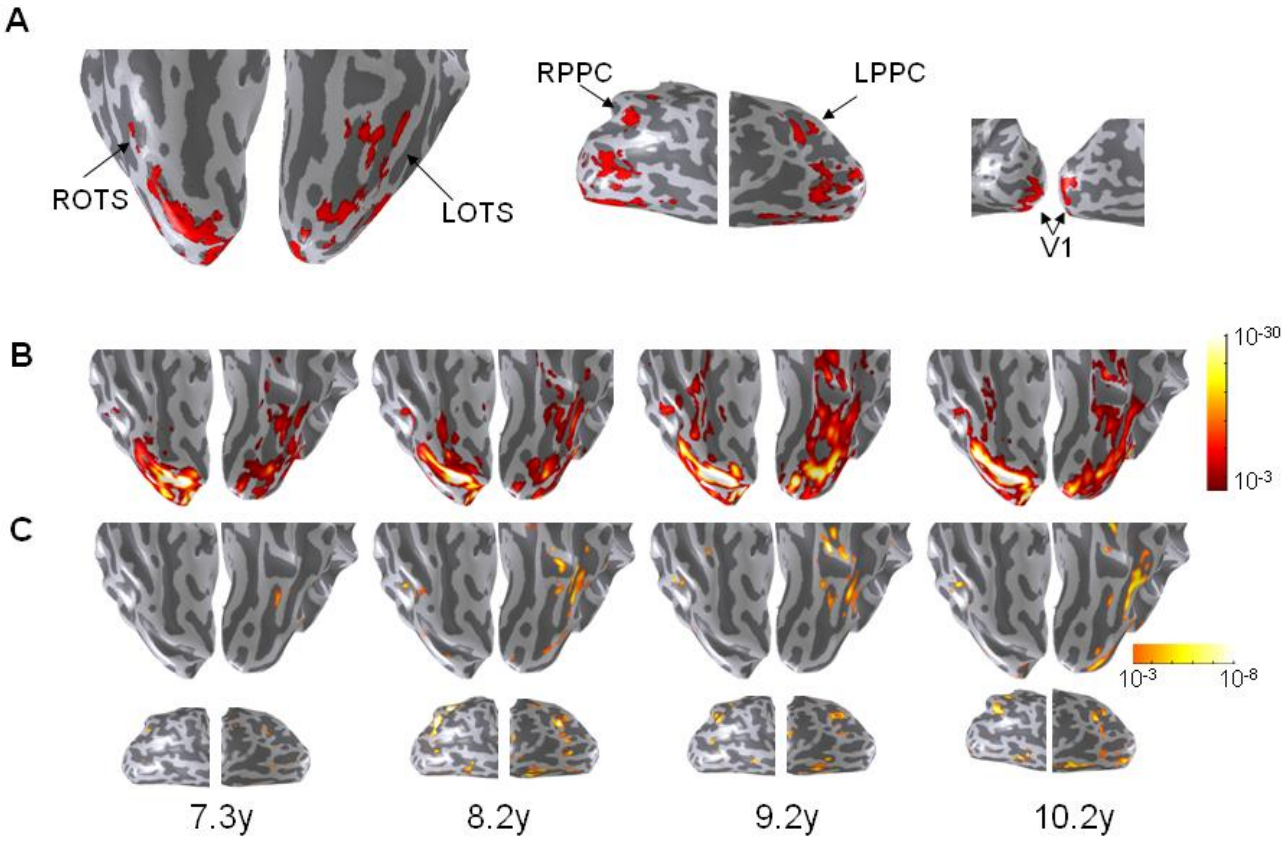

Figure S6. Individual contrast maps across 4 measurement points. S12, female, entry age 7.3y. A. ROI definition. LOTS is shown on a ventral view, PPC is shown on a lateral view, V1 is shown on a medial view. B. Localizer contrast map over 4 measurement points. C. word visibility contrast maps. Threshold is $p<0.001$, uncorrected.

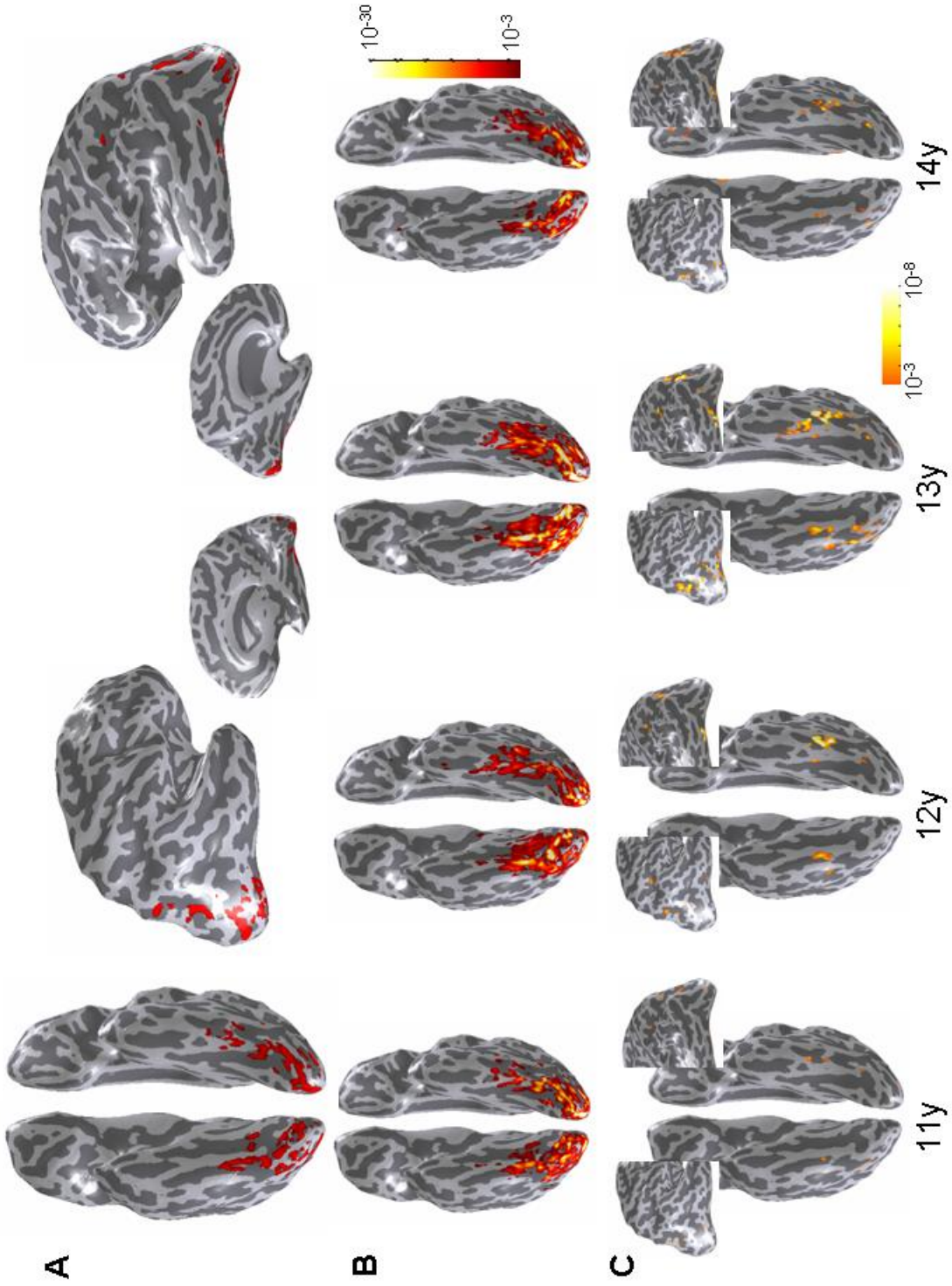

Figure S7. Individual contrast maps across 4 measurement points. S8, female, entry age 11y, poor reader. See caption to Figure S6 for more detail.

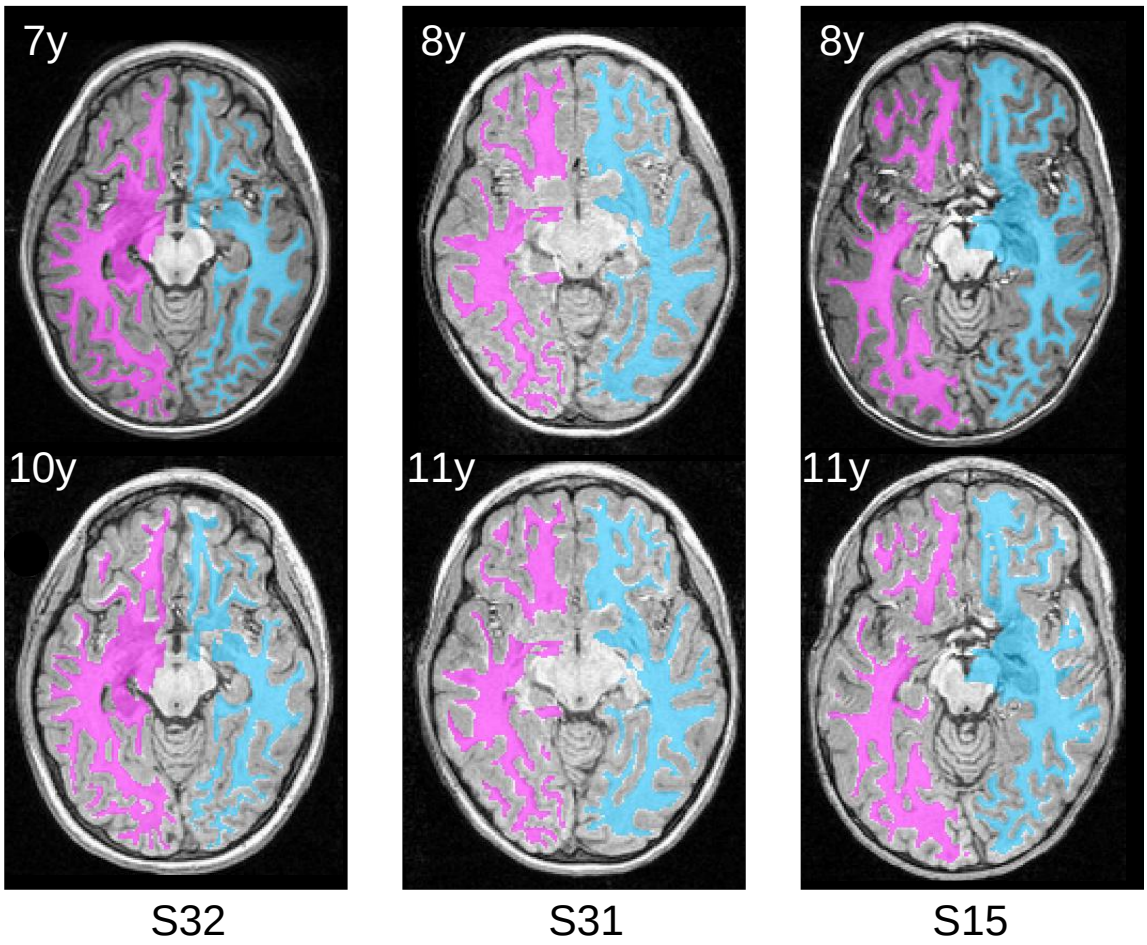

Figure S8. Anatomical change in individuals across 3 years. Tissue segmentation of the anatomical images acquired in the first measurement (top row) is overlaid on the T1 anatomical images taken 3 years later (bottom row) in three individuals who entered the study at a young age. White matter of the left and right hemispheres are shown in magenta and cyan. The axial slices are at the approximate level of the LOTS. There are minimal anatomical changes around the posterior part of the brain, though more extensive changes (noncolored white matter) appear in frontal and anterior temporal regions. Taking into account the spatial resolution of functional measurements (2.5x2.5x3mm voxels, compare with $1mm^3$ here), these images suggest that changes in VWFA neurometric functions across the years are unlikely to stem from sampling different parts of cortex in different years. Rather, the changes are likely to represent development in the functional profile of this region over time.

Appendix I: Word signals along the ventral surface

Several recent studies propose that word selectivity emerges as we move anteriorly along the ventral occipito-temporal surface (e.g., Brem et al., 2006, 2009; Vinckier et al., 2007). Our study does not test selectivity but it allows a crude assessment of age related and reading skill related differences in word responses along the posterior-anterior gradient. For this purpose, we implemented the analysis from Brem et al., 2006, and compared word signals within MNI spheres planted along the ventral surface.

Analysis I consists of a comparison of word responses across age groups and between good and poor readers.

Analysis II is a refinement of analysis I, restricting the MNI defined spheres to gray matter voxels. This analysis is expected to enhance the signals because it excludes white matter voxels which contribute mostly noise to the mean amplitudes.

Analysis III takes a data driven approach and examines potential changes in the location of the visibility effect over time. If this analysis finds a significant shift of active voxels anteriorly, it could cast doubts on our attempt to analyze changes in sensitivity from a fixed set of voxels over time (main paper, figures 1-4).

Analysis I: Word responses along the posterior-anterior axis

We defined regions of interest along the posterior-anterior axis based on the prescription in Brem et al., 2006, 2009. Five spheres were defined, 6mm radius each, around the MNI coordinates specified in these papers. We compared % signal change for fully visible words (vs. fixation) in each sphere, for the three age groups (Figure S9). Younger children show slightly stronger signals in the posterior region (sphere 1), and weaker responses in the most anterior region (sphere 5). These differences, though weak, are consistent with some developmental trend towards stronger anterior word responses and inhibition of word responses in posterior VOT.

A comparison of good and poor readers' word responses within the same spheres did not reveal any differences, as shown in Figure S10.

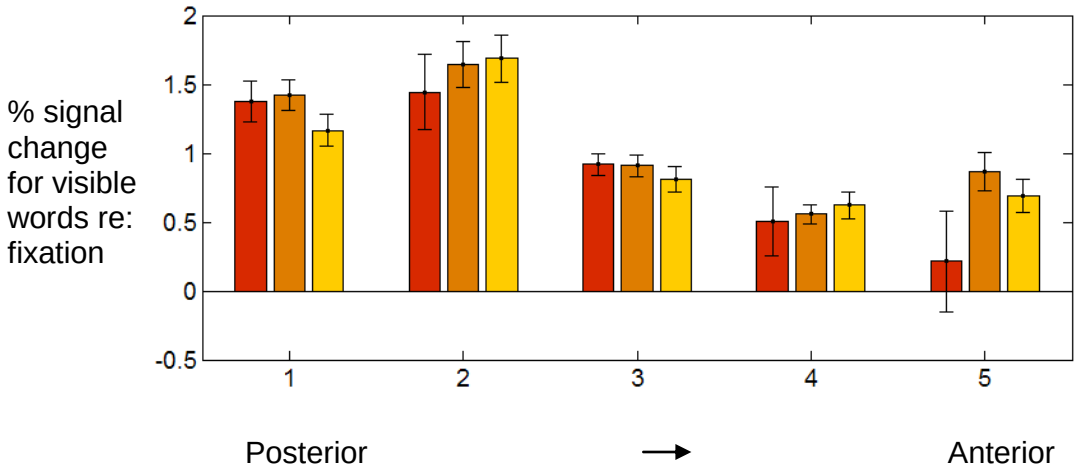

Figure S9. Age grouped fMRI word-signals along the posterior anterior axis. FMRI responses are shown for the visible words condition only, within 5 spheres along the posterior-anterior axis. Signals are compared between 3 age groups: 7-8y (red), 9-11y (orange), 12-15y (yellow).

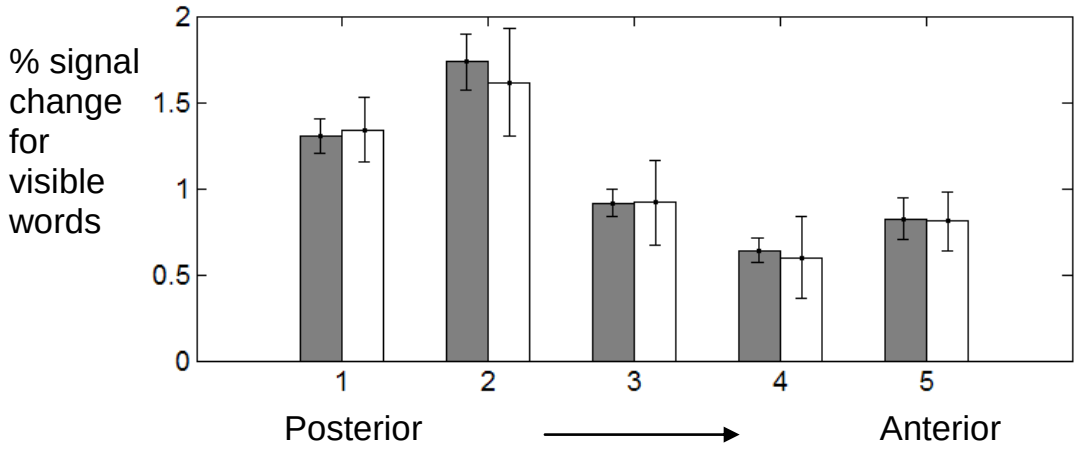

Figure S10. Grouping by reading skill. fMRI responses are shown for the visible words condition only, within 5 spheres along the posterior-anterior axis. Signals are compared between good readers (gray, basic reading age standardized score >=100) and poor readers (white, basic reading age standardized score <90).

Analysis II: Refined analysis of word responses along the ventral surface

MNI based spheres include many voxels that are irrelevant to the measured signals, and these are likely to contribute noise to this analysis. For example, some of the voxels included in these spheres are placed in white matter, which does not contribute interpretable BOLD signals. We refined analysis I by restricting the MNI defined spheres to gray matter voxels (based on gray-white segmentations of the anatomical images). We

used FSL automatic segmentation tool (FAST) to segment the high resolution anatomical MRIs, and restricted the analysis to gray matter voxels within each sphere.

Figure S11-12 parallel Fig. S9-10 with restricted spheres.

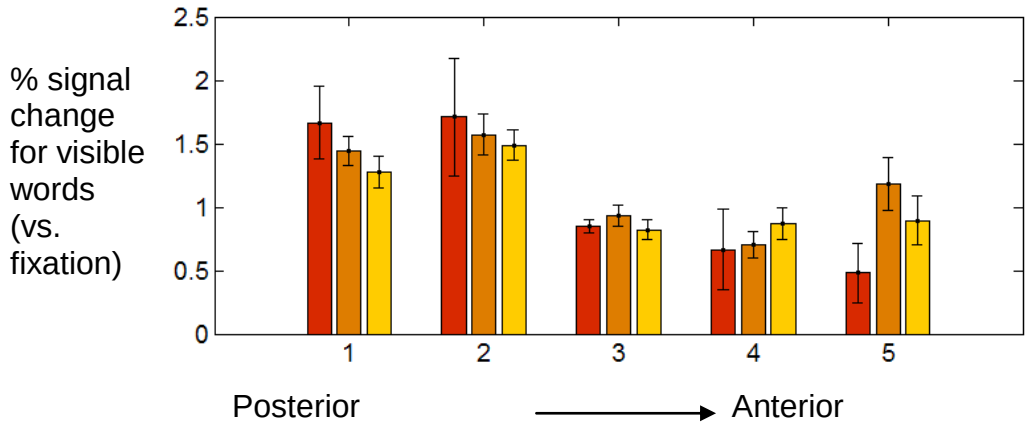

Figure S11. Age grouped fMRI word-signals along the posterior anterior axis. fMRI responses are shown for the visible words condition only, in gray matter voxels only within 5 6mm radius spheres along the posterior-anterior axis. Signals are compared between 3 age groups: 7-8y (red), 9-11y (orange), 12-15y (yellow).

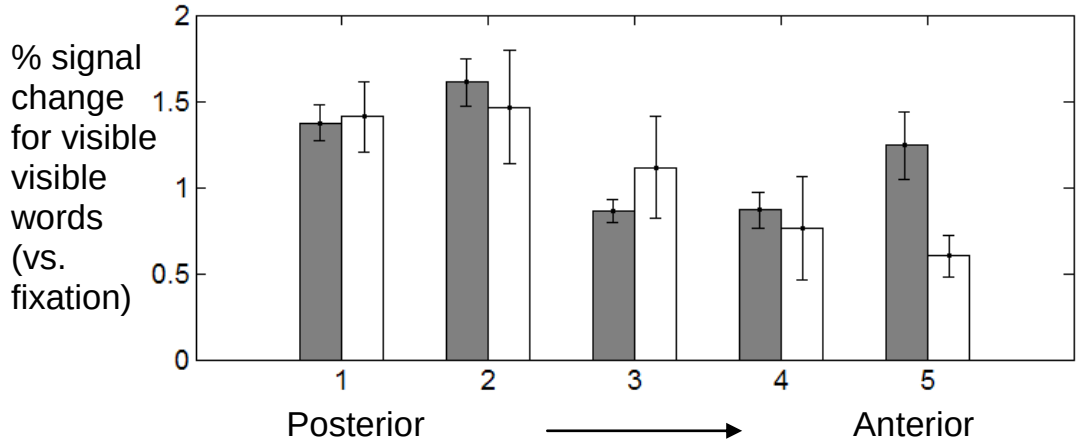

Figure S12. Grouping by reading skill. fMRI responses are shown for the visible words condition only, **in gray matter voxels only** within 5 spheres along the posterior-anterior axis. Signals are compared between good readers (gray, basic reading age standardized score >=100) and poor readers (white, basic reading age standardized score <90).

As shown in Fig. S11, the general pattern of age differences remains: Younger children show slightly stronger signals in the posterior region (sphere 1) and weaker responses in the most anterior region (sphere 5) compared to older children.

For the grouping by reading skill, restricting the spheres to gray matter voxels had a larger effect. Specifically, the refined analysis reveals a clear effect in the most anterior

regions, with poor reader showing much weaker signals. There is also a slightly higher word activation in the poor readers in sphere 3, echoing our finding in the functionally defined region (See figure 4 in the main paper).

Analysis III: Do the word visibility activations move forward with age?

Our longitudinal analysis localizes voxels that are consistently activated by our functional localizer (all word conditions vs fixation) across the years, and examines changes in the neurometric curves within these voxels.

It is possible, however, that the spatial distribution of the word visibility responses is changing over time, and such a change could be missed by focusing on the same region over time. To assess whether the word visibility responses are moving forward over time, we identified (for each child for each time point) left and right occipito-temporal clusters activated by the word visibility contrast (2 more visible conditions vs. two less visible conditions; $p<0.001$ uncorrected). We then examined changes and differences in the center of mass of this region.

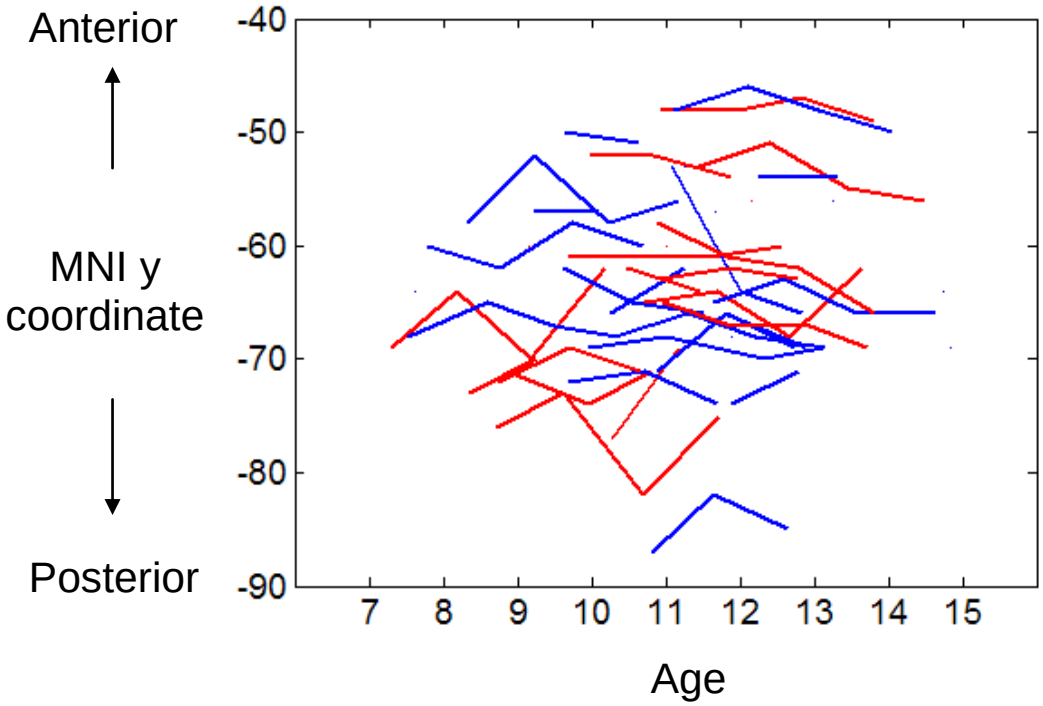

Figure S13. Individual "growth curves" of the MNI y coordinate of the cluster activated by the visibility contrast. Curves are colored for better readers (basic reading standardized score >=100, blue) and worse readers (BRss<100, red).

As shown in Figure S13, there is no clear shift of the center of mass of activation towards the anterior direction (larger y coordinates). For each child, the location of the word visibility activations remains stable along the anterior-posterior axis. The same is true for the left-right and inferior-superior axes (data not shown). We therefore conclude that our analysis of sensitivity change within the same region over time is not hindered by changes in the spatial distribution of the signals.