Check for updates

Cortical recycling in high-level visual cortex during childhood development

Marisa Nordt [1], Jesse Gomez [2], Vaidehi S. Natu[1], Alex A. Rezai [1], Dawn Finzi [1], Holly Kular[1] and Kalanit Grill-Spector [1,3,4] ✉

Human ventral temporal cortex contains category-selective regions that respond preferentially to ecologically relevant categories such as faces, bodies, places and words and that are causally involved in the perception of these categories. How do these regions develop during childhood? We used functional magnetic resonance imaging to measure longitudinal development of category selectivity in school-age children over 1 to 5 years. We discovered that, from young childhood to the teens, face- and word-selective regions in ventral temporal cortex expand and become more category selective, but limb-selective regions shrink and lose their preference for limbs. Critically, as a child develops, increases in face and word selectivity are directly linked to decreases in limb selectivity, revealing that during childhood, limb selectivity in ventral temporal cortex is repurposed into word and face selectivity. These data provide evidence for cortical recycling during childhood development. This has important implications for understanding typical as well as atypical brain development and necessitates a rethinking of how cortical function develops during childhood.

A central question in neuroscience is how cortical function develops. Ventral temporal cortex (VTC) is an excellent model system to address this question as it contains regions selective for ecological categories such as faces[1], bodies[2], places[3] and words[4] that are critical for human behaviour and can be identified in each individual. When infants first open their eyes, they are inundated with faces, body parts, their surrounding room, and objects. This visual input may begin to shape VTC representations in infancy and lead to the emergence of face representations in the first year of life[5–7]. However, experience with other categories, such as words, does not begin until later in childhood when children learn to read.

Two main theories regarding the development of category-selective regions have been proposed. The theory of functional refinement predicts that category-selective regions emerge from raw, general-purpose cortex[8,9], which has some basic property[10] that is present early in development such as foveal bias[11,12]. For example, a longitudinal study that tracked development of word selectivity during the first year of school found that word selectivity emerged upon previously unspecialized cortex[9]. Similarly, developmental studies of face-selective regions reported that the growing part of face-selective regions showed less specificity in children than adults[8] and also in younger than older infant monkeys[7].

Alternatively, the theory of competition posits that rivalry for cortical resources may constrain development[13,14]. In particular, as both face and word recognition require fine visual acuity afforded by foveal vision, emerging representations for these categories may compete for foveal resources in lateral VTC[15]. This competition for foveal cortical territory may lead to recycling of cortex selective to one category earlier in childhood (for example, faces) to be selective to another category (for example, words) with later demands for reading[14,16]. Indeed, comparison of brains of illiterate or late-literate adults with literate adults suggests that increased literacy in adulthood is associated with reduced responses to faces in the left

fusiform gyrus[14,17]. Because reading acquisition is specific to humans and because the fusiform gyrus, which is the anatomical structure where face-selective regions reside, is hominoid specific, addressing this intense theoretical debate necessitates longitudinal brain measurements in school-age children to examine the development of multiple category representations.

Here, we addressed this key gap in knowledge using longitudinal functional magnetic resonance imaging (fMRI) in 29 children (initially 5–12 years old) to measure development of cortical representations of multiple categories as well as their relationship. Children were scanned repeatedly over the course of 1 to 5 years (mean ± s.d. 3.75 ± 1.5 years; Supplementary Fig. 1A) with an average of 4.4 ± 1.92 fMRI sessions per child and 128 included sessions overall (Methods). During fMRI, children viewed images from ten categories spanning five domains: characters (pseudowords, numbers), body parts (headless bodies, limbs), faces (adult faces, child faces), objects (string instruments, cars) and places (houses, corridors) (Supplementary Fig. 1D). Analyses were performed in each individual's native brain space, which allows precise tracking of the developing cortex in each participant and prevents blurring of responses from different categories due to normalization to a standard brain template[18].

Results

How does category selectivity develop in VTC? To assess the development of category selectivity in VTC, we first quantified the volume of category-selective activations in VTC as a function of age. Selectivity was computed by contrasting responses to each category versus all other categories except the other category from the same domain (for example, limbs versus all other categories except bodies, $t > 3$, voxel level). VTC was defined anatomically on the cortical surfaces of each child and divided into lateral and medial partitions (Fig. 1b). This division captures the centre–periphery organization of VTC[12,15], where lateral VTC represents the central visual field

[1]Department of Psychology, Stanford University, Stanford, CA, USA. [2]Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA. [3]Neurosciences Program, Stanford University, Stanford, CA, USA. [4]Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, USA. ✉e-mail: kalanit@stanford.edu
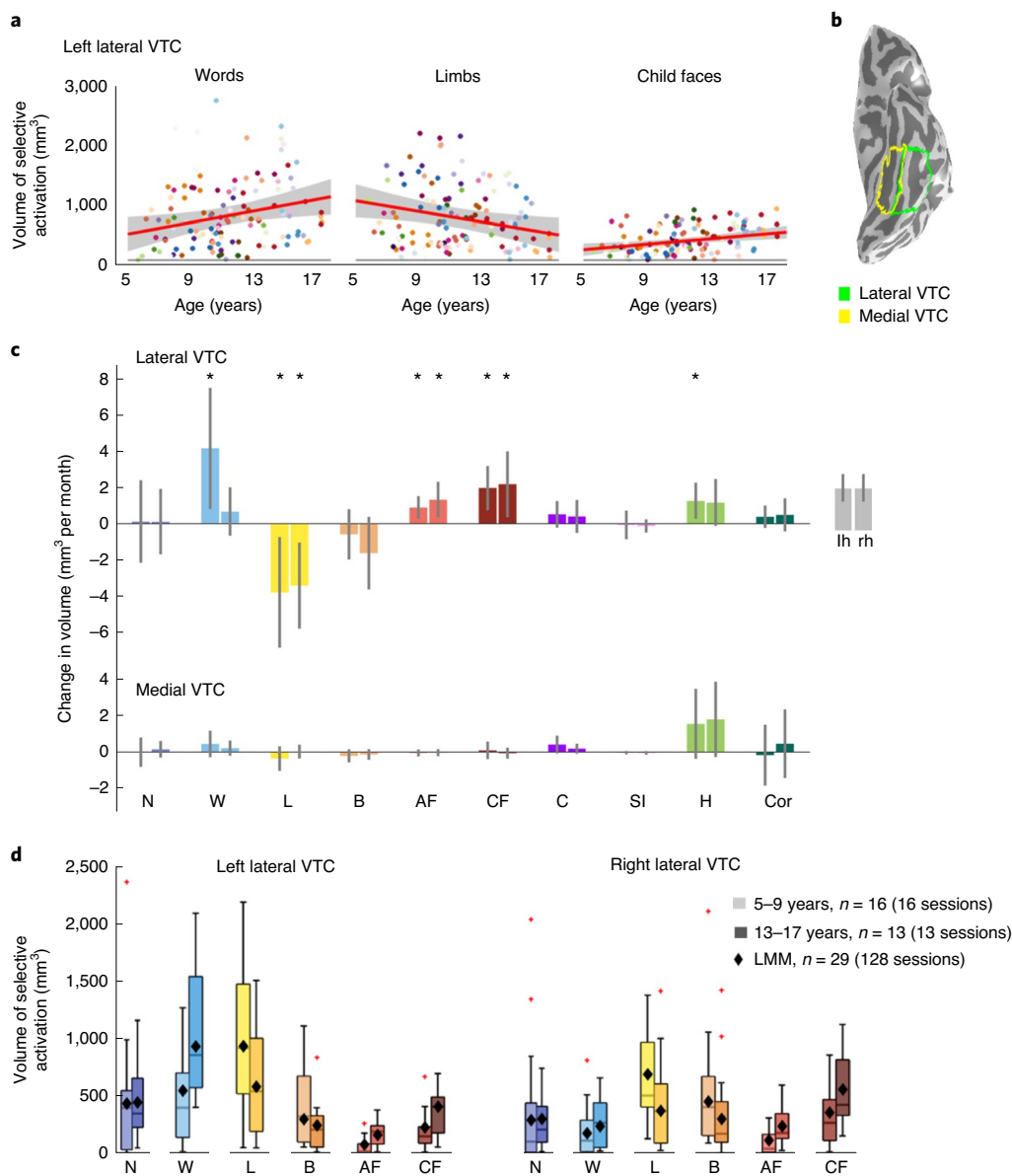
Fig. 1 | Developmental increases and decreases in category-selective activation in lateral VTC. a, Volume of word-, limb- and child-face-selective activation by age in each session (dot, coloured by participant), plus LMM prediction of category-selective activation by age (red line) and 95% CI (grey shading). All scatterplots in Supplementary Fig. 2. **b**, Lateral and medial VTC on the inflated cortical surface of a 5-year-old. **c**, LMM slopes indicating change in category-selective activation volume per month ($n = 128$ sessions, 29 children) with 95% CI. If the CI does not cross the $y = 0$ line, the slope is significantly different from 0 before FDR correction. Asterisks indicate significant ($P < 0.05$) development after FDR correction. **d**, Category-selective activation by age group, showing average category-selective volume for 5–9- and 13–17-year-olds (boxplots, with 25% to 75% percentiles coloured, median as horizontal line and whiskers to extremes not considered outliers, which are marked with crosses for values more than 1.5 times the interquartile range away from the bottom or top of the box) with one session per child (number of sessions in legend) and LMM estimate for category-selective volume (diamonds) for the mean age of each age group. LMMs are based on data from all sessions, as in **a** and **c**. All categories in Supplementary Fig. 3. N, numbers; W, words; L, limbs; B, bodies; AF, adult faces; CF, child faces; C, cars; SI, string instruments, H, houses; Cor, corridors.

and medial VTC represents the peripheral visual field[19]. To test for age-related changes in the volume of category-selective activations, we used linear mixed models (LMMs) with age as a fixed effect and participant as a random effect. LMMs with intercepts that varied across participants and a constant slope fitted the data best in the majority of cases (Methods). Figure 1a shows examples of LMM fits for words, limbs and faces (all LMMs in Supplementary Fig. 2). Figure 1c summarizes the LMM slopes and confidence intervals (CI) for all categories and VTC partitions. The slopes (β_{age}) indicate the rate of change (mm³ per month) of category-selective volume

with age (Fig. 1c). Significant development after false discovery rate (FDR) correction[20] is indicated by asterisks in Fig. 1c (full statistics in Supplementary Tables 1 and 2).

Results reveal variable development of category selectivity across VTC partitions and categories. Changes in category selectivity were significant in lateral VTC but not statistically significant in medial VTC (Fig. 1c, Supplementary Fig. 2 and Supplementary Tables 1 and 2). Surprisingly, there were both developmental increases and decreases in the volume of category-selective activation in lateral VTC: word- and face-selective activation significantly increased,

but limb-selective activation significantly decreased (Fig. 1a,c and Supplementary Fig. 2).

Interestingly, during childhood, the volume of word-selective activation significantly increased in left lateral VTC (Fig. 1c, light blue; $\beta_{age} = 4.14$ mm³ per month, 95% CI 0.81–7.48 mm³ per month, $t(126) = 2.46$, $P_{FDR} = 0.044$). However, there was no statistically significant change in word-selective activation in right lateral VTC (Fig. 1c,d and Supplementary Fig. 2; $\beta_{age} = 0.66$ mm³ per month, 95% CI −0.66 to 1.99 mm³ per month, $t(126) = 0.99$, $P_{FDR} = 0.46$) or of number-selective activations bilaterally (Fig. 1c, Supplementary Fig. 2 and Supplementary Table 1; left: $\beta_{age} = 0.11$ mm³ per month, 95% CI −2.15 to 2.38 mm³ per month, $t(126) = 0.10$, $P_{FDR} = 0.92$; right: $\beta_{age} = 0.11$ mm³ per month, 95% CI −1.69 to 1.90 mm³ per month, $t(126) = 0.12$, $P_{FDR} = 0.92$). Examining the average measured volume of activations (Fig. 1d, boxplots; Supplementary Fig. 3 for all categories) as well as LMM predictions of the mean volume (Fig. 1d, diamonds) revealed that word selectivity doubled on average from ~500 mm³ in 5–9-year-olds to ~1,000 mm³ in 13–17-year-olds. Notably, as the volume of word-selective activation doubled, the volume of limb-selective activation halved (Fig. 1d, yellow). In fact, limb-selective activation significantly decreased over childhood in both hemispheres (Fig. 1c, yellow; left: $\beta_{age} = -3.79$ mm³ per month, 95% CI −6.84 to −0.74 mm³ per month, $t(126) = -2.46$, $P_{FDR} = 0.044$; right: $\beta_{age} = -3.42$ mm³ per month, 95% CI −5.79 to −1.04 mm³ per month, $t(126) = -2.85$, $P_{FDR} = 0.04$). However, there was no statistically significant change in body-selective activation (Fig. 1c, orange; left: $\beta_{age} = -0.59$ mm³/month, 95% CI −1.98 to 0.80 mm³ per month, $t(126) = -0.84$, $P_{FDR} = 0.50$; right: $\beta_{age} = -1.63$ mm³ per month, 95% CI −3.63 to 0.38 mm³ per month, $t(126) = -1.61$, $P_{FDR} = 0.22$). At the same time, the volume of face-selective activation significantly increased over childhood bilaterally for both adult faces (Fig. 1c, bright red; left: $\beta_{age} = 0.89$ mm³ per month, 95% CI 0.26–1.52 mm³ per month, $t(126) = 2.79$, $P_{FDR} = 0.04$; right: $\beta_{age} = 1.32$ mm³ per month, 95% CI 0.34–2.29 mm³ per month, $t(126) = 2.68$, $P_{FDR} = 0.042$) and child faces (Fig. 1c, dark red; left: $\beta_{age} = 1.95$ mm³ per month, 95% CI 0.74–3.17 mm³ per month, $t(126) = 3.18$, $P_{FDR} = 0.037$; right: $\beta_{age} = 2.16$ mm³ per month, 95% CI 0.35–3.97 mm³ per month, $t(126) = 2.37$, $P_{FDR} = 0.049$). Similar longitudinal development in lateral VTC was observed for other contrasts, such as contrasting responses by domain (Supplementary Fig. 4) and for other metrics, such as the magnitude of category selectivity within constant-sized regions (Supplementary Fig. 5).

As limb representations are located between face and word activations, and the latter show developmental increases, it is possible that the developmental decrease in volume of limb-selective activation may result from the inclusion of faces and words as control categories when defining category contrasts. To test this possibility, we repeated the analyses in Fig. 1 but excluding faces, words and limbs from the control categories in the contrast. Results of these analyses reveal a qualitative similar pattern of results, although they do not survive FDR correction (Extended Data Fig. 1a,c, Supplementary Fig. 6 and Supplementary Tables 3 and 4). That is, word-selective volume increased in the left hemisphere, face-selective volume increased bilaterally while limb-selective volume decreased bilaterally with childhood development.

Another concern is that age-related differences in data quality may conflate developmental effects. Thus, we performed several controls to test whether factors that may affect data quality contribute to these findings. First, behavioural performance on the oddball task performed during scanning was overall high (median performance 91%, s.d. 18%), indicating that children were attending to the stimuli. Second, to test whether developmental effects may be related to temporal signal to noise (tSNR) or motion during scan, we ran new LMMs which included, in addition to age, tSNR and average motion during scan as predictors. Motion was not a significant predictor of the volume of category-selective activation

(Supplementary Table 5). While tSNR was a significant predictor, it was independent from age. That is, LMMs that included tSNR as an additional predictor revealed the same pattern of results showing significant age-related increases in face- and word-selective activations as well as significant age-related decreases in limb-selective activations (Extended Data Fig. 1b,c and Supplementary Tables 6 and 7). Third, to test whether developmental effects are driven by overall lower reliability of responses in younger children compared with older children, we assessed response reliability in V1 as a control region of interest (ROI). Response reliability in V1 was numerically high (left: median 0.74, s.d. 0.14; right: median 0.56, s.d. 0.16) and was not statistically significant related to age (Supplementary Fig. 7). Together, these analyses validate that age-related changes in category-selective activations reflect longitudinal development rather than differences in measurement quality across years.

Is development anatomically specific? To determine the anatomical specificity of the observed development, we defined limb-, word- and face-selective ROIs in each participant and session (Fig. 2a–c; $t > 3$ voxel level) and examined them longitudinally. Category-selective ROIs remained largely within the same anatomical structures as ROIs expanded or shrank across childhood (Fig. 2a–c). We used LMMs to examine the effect of age on ROI volume (age as fixed effect, participant as random effect; change of ROI volume per month (β_{age}) in Fig. 2d, full statistics in Supplementary Table 8).

We found that posterior ROIs significantly expanded with age. In contrast, the volume of anterior ROIs did not significantly change with age (Fig. 2d). Activation for words grew significantly in the left posterior occipitotemporal sulcus (Fig. 2d, pOTS-words; $\beta_{age} = 1.71$ mm³ per month, 95% CI 0.45–2.98 mm³ per month, $t(117) = 2.68$, $P_{FDR} = 0.02$), but there was no statistically significant effect of age on activation in the mid occipitotemporal sulcus (Fig. 2d, mOTS-words; left: $\beta_{age} = -0.07$ mm³ per month, 95% CI −1.22 to 1.08 mm³ per month, $t(101) = -0.12$, $P_{FDR} = 0.90$; right: $\beta_{age} = 0.25$ mm³ per month, 95% CI −0.43 to 0.92 mm³ per month, $t(33) = 0.74$, $P_{FDR} = 0.66$). Similarly, bilateral activation for faces grew significantly in the posterior fusiform (Fig. 2d, pFus-faces; left: $\beta_{age} = 0.75$ mm³ per month, 95% CI 0.17–1.33 mm³ per month, $t(118) = 2.55$, $P_{FDR} = 0.02$; right: $\beta_{age} = 1.41$ mm³ per month, 95% CI 0.46–2.35 mm³ per month, $t(96) = 2.96$, $P_{FDR} = 0.019$). However, there was no statistically significant effect of age on activation in the mid-fusiform (Fig. 2d, mFus-faces; left: $\beta_{age} = 0.11$ mm³ per month, 95% CI −0.47 to 0.69 mm³ per month, $t(105) = 0.38$, $P_{FDR} = 0.79$; right: $\beta_{age} = 0.20$ mm³ per month, 95% CI −0.46 to 0.85 mm³ per month, $t(100) = 0.60$, $P_{FDR} = 0.69$). In contrast to the growth of pOTS-words and pFus-faces, OTS-limbs shrank significantly in both hemispheres (Fig. 2d, OTS-limbs and Supplementary Table 8; left: $\beta_{age} = -1.70$ mm³ per month, 95% CI −2.59 to −0.81 mm³ per month, $t(124) = -3.78$, $P_{FDR} = 0.002$; right: $\beta_{age} = -0.89$ mm³ per month, 95% CI −1.53 to −0.24 mm³ per month, $t(124) = -2.73$, $P_{FDR} = 0.02$). Control analyses showed that (i) the surface area of face- and word-selective regions significantly expanded and the surface area of limb-selective regions significantly shrank (Supplementary Fig. 8) and (ii) the expansion of face-selective regions and the shrinking of limb-selective regions were also significant for ROIs defined from domain contrasts (Supplementary Fig. 9).

What functional changes occur in the developing regions? We next assessed the functional properties of the developing regions of category-selective ROIs in each participant. For words and faces, where selective voxels emerge during development, we call the difference between the end and initial ROIs the emerging pOTS-words and emerging pFus-faces. For limbs, we call the difference between the initial and end ROIs the waning OTS-limbs. We focus on the left hemisphere due to the observed left lateralized development
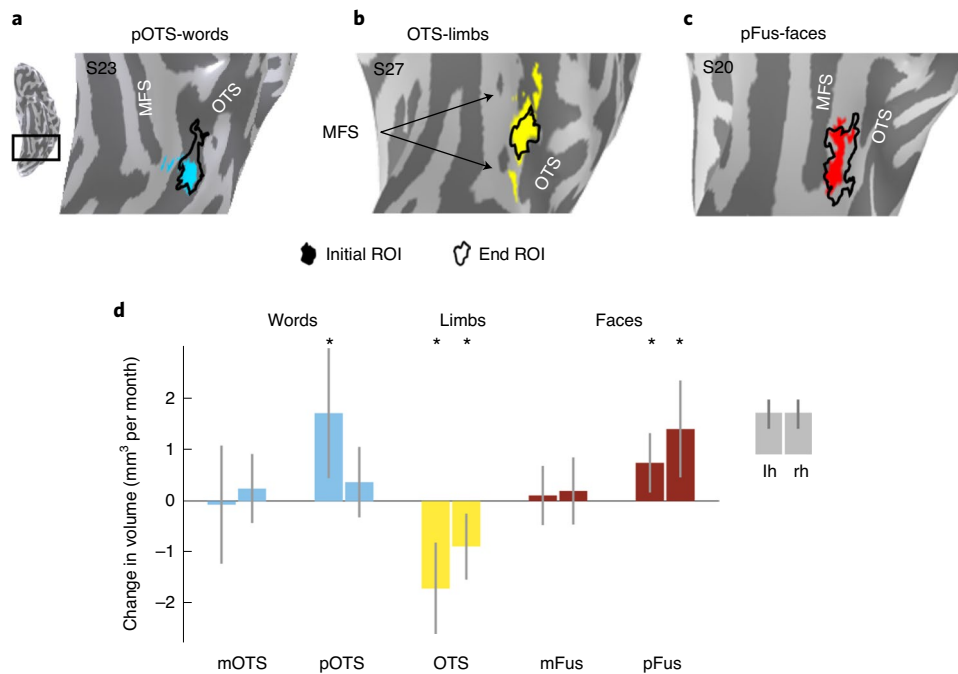
Fig. 2 | Development of category-selective ROIs. a–c, Initial ROIs (coloured) and end ROIs (outline) in three example children for left pOTS-words at age 10 and 15 years (**a**), left OTS-limbs at age 11 and 13 years (**b**) and left pFus-faces at age 9 and 14 years (**c**). **d**, LMM slopes showing change in volume of category-selective regions per month with 95% CI (error bars). If the CI does not cross the $y=0$ line, the slope is significantly different from 0 before FDR correction. Asterisks indicate significant development ($P<0.05$, FDR corrected). Number of sessions per ROI is $n=103$ for left mOTS-words, $n=35$ for right mOTS-words, $n=119$ for left pOTS-words, $n=73$ for right pOTS-words, $n=126$ for left OTS-limbs, $n=126$ for right OTS-limbs, $n=107$ for left mFus-faces, $n=102$ for right mFus-faces, $n=120$ for left pFus-faces and $n=98$ for right pFus-faces. MFS, mid-fusiform sulcus; lh, left hemisphere; rh, right hemisphere; OTS, occipitotemporal sulcus; S23, S27 and S20, participants 23, 27 and 20, respectively.
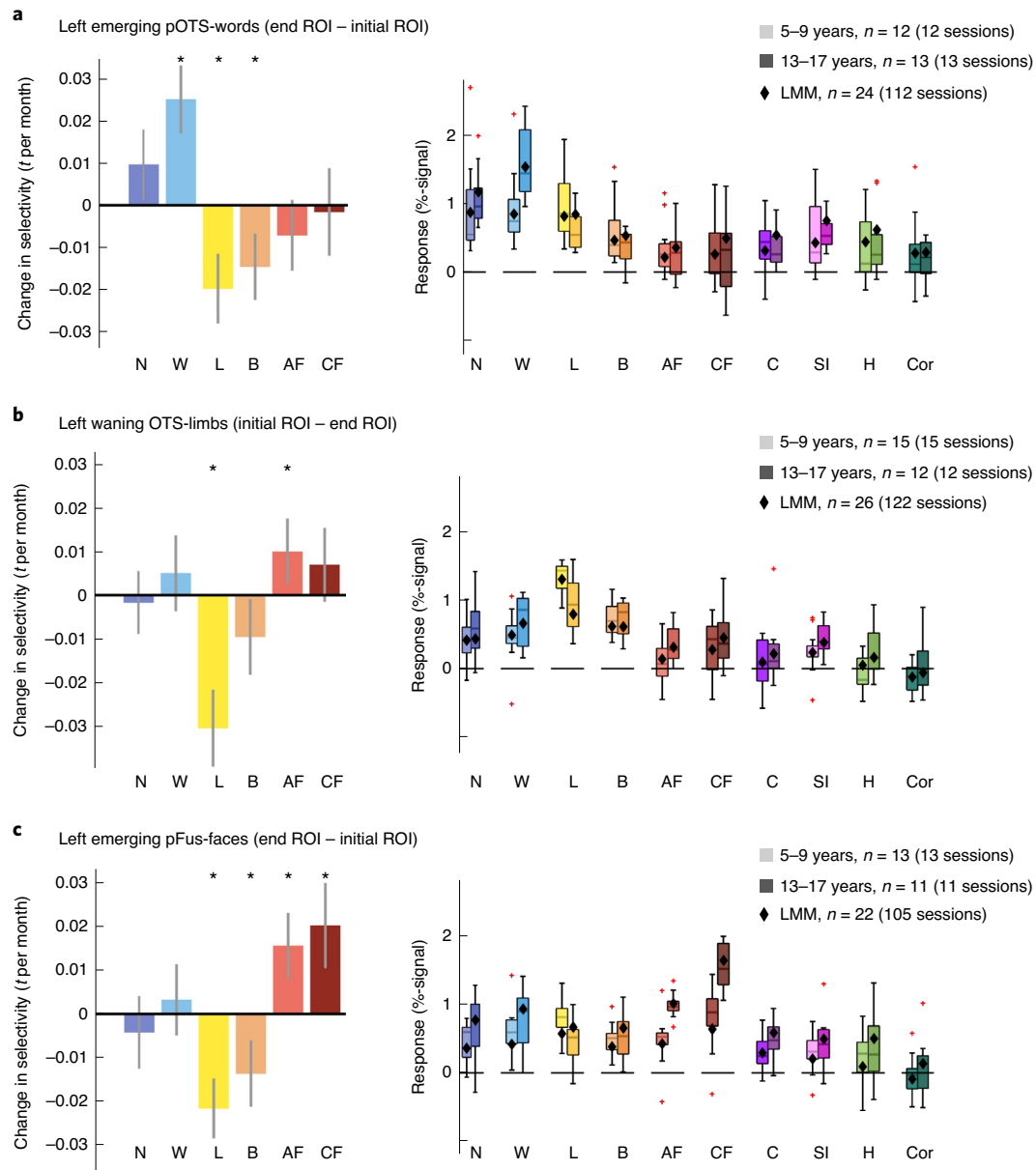
of word selectivity (right hemisphere in Extended Data Fig. 2). LMMs were used to assess the effect of age on selectivity in emerging and waning ROIs (age as continuous fixed effect, participant as random effect). LMM slopes (change in selectivity, t per month) and their significance are shown in Fig. 3 (left; statistics in figure caption and Supplementary Table 9). Since developing regions are not completely independent from the original ROIs, we repeated the analysis in independent ring-shaped ROIs centred on the initial functional ROIs, yielding similar findings (Supplementary Fig. 10).

In the emerging left pOTS-words, word selectivity significantly increased with age, as expected from the definition of the ROI (Fig. 3a, left, blue bar). At the same time, limb selectivity and body selectivity significantly decreased (Fig. 3a, left), but we found no statistically significant changes in selectivity to other categories (Extended Data Fig. 2 and Supplementary Table 9). As a control, we repeated these analyses but excluding adult faces, child faces, limbs and words as control categories from all contrasts, finding similar results (Fig. 3A, left, Extended Data Fig. 2, open bars and Supplementary Table 10).

To elucidate whether the changes in selectivity in the emerging ROI were due to increased responses to the preferred category or decreased responses to nonpreferred categories, we used LMMs to quantify the response amplitude to each category as a function of age (age as continuous fixed effect, participant as random effect, change in %-signal per month (β_{age}); Supplementary Fig. 11; statistics in Supplementary Table 11). Results show that developmental changes in word selectivity were associated with significant increases in the responses to words ($\beta_{age}=0.008$ %-signal per month, 95% CI 0.004–0.012 %-signal per month, $t(110)=3.82$, $P_{FDR}=0.002$) with no statistically significant changes in responses to other categories (Supplementary Fig. 11 and Supplementary Table 11) except

that responses to string instruments also significantly increased ($\beta_{age}=0.004$ %-signal/month, 95% CI 0.001–0.006 %-signal per month, $t(110)=2.98$, $P_{FDR}=0.025$). Indeed, average responses to words in pOTS-words are higher in teens (13–17-year-olds) than children (5–9-year-olds) (Fig. 3a, right, boxplots, Supplementary Fig. 11 and Supplementary Table 11) in correspondence with LMM predictions (Fig. 3a, right, diamonds).

Similarly, in the emerging pFus-faces, selectivity to faces increased (Fig. 3c, left, red bars and Extended Data Fig. 2). At the same time, selectivity to limbs significantly decreased bilaterally (Fig. 3c, left, yellow bar), and this development was significant in the left hemisphere also when faces and words were excluded as control categories from the contrast (Fig. 3c, left and Extended Data Fig. 2, open bars); selectivity to bodies decreased significantly only in the left hemisphere (Fig. 3c, orange). We found no statistically significant changes in selectivity to other categories (Extended Data Fig. 2 and Supplementary Table 9). Increases in selectivity to faces were associated with significant increases in responses to faces (Supplementary Fig. 11 and Supplementary Table 11; adult faces: $\beta_{age}=0.006$ %-signal/month, 95% CI 0.004–0.009 %-signal/month, $t(103)=5.33$, $P_{FDR}<0.001$; child faces: $\beta_{age}=0.011$ %-signal/month, 95% CI 0.007–0.015 %-signal/month, $t(103)=5.79$, $P_{FDR}<0.001$). Indeed, average responses to both adult and child faces were higher in teens than children, in correspondence with LMM predictions (Fig. 3c, right). Additionally, responses to words and string instruments also significantly increased in the left emerging pFus-faces (words: $\beta_{age}=0.006$ %-signal/month, 95% CI 0.002–0.01 %-signal/month, $t(103)=2.78$, $P_{FDR}=0.04$; string instruments: $\beta_{age}=0.003$ %-signal/month, 95% CI 7.7×10^{-4} to 0.006 %-signal/month, $t(103)=2.62$, $P_{FDR}=0.04$), and there was a trend for an increase in responses to cars (Fig. 3c, right, Supplementary Fig. 11 and Supplementary

a Left emerging pOTS-words (end ROI – initial ROI)

b Left waning OTS-limbs (initial ROI – end ROI)

c Left emerging pFus-faces (end ROI – initial ROI)


Fig. 3 | Age-related increases in word and face selectivities parallel decreases in limb selectivity in the developing regions. a–c, Left: LMM slopes indicating changes in selectivity by age with 95% CI (error bars). If the CI does not cross the $y = 0$ line, the slope is significantly different from 0 before FDR correction. Asterisks indicate significant development after FDR correction ($P < 0.05$). All ROIs and categories in Extended Data Fig. 2. Full statistics in Supplementary Tables 9 and 10. Right: response amplitudes for the ten categories shown as boxplots illustrating response amplitudes by age group for 5–9-year-olds and 13–17-year-olds with one session per child (legend indicates number of sessions). LMMs on response amplitudes using data from all available sessions for each ROI were used to estimate development of response amplitudes. Diamonds represent LMM estimates for response amplitudes for the mean age of each age group. Full statistics in Supplementary Table 11. Boxplots show the 25% to 75% percentiles (coloured areas) and median (horizontal lines). Whiskers extend to the most extreme data points not considered outliers (minimum, maximum). Crosses show outliers (values more than 1.5 times the interquartile range away from the bottom or top of the box).

Table 11). Thus, developmental increases in word and face selectivity in emerging word and face ROIs, respectively, are largely driven by increased response amplitudes to the preferred category rather than decreased response amplitudes to nonpreferred categories.

As OTS-limbs is located between pOTS-words and pFus-faces, we asked whether increased responses to faces and words also occur in waning OTS-limbs. We found that, not only did responses to adult faces significantly increase with age (Supplementary Fig. 11; $\beta_{age} = 0.002$ %-signal/month, 95% CI 4.6×10^{-4} to 0.003 %-signal/month, $t(120) = 2.64$, $P_{FDR} = 0.043$), but also responses to limbs

significantly decrease with age (Supplementary Fig. 11; $\beta_{age} = -0.005$ %-signal/month, 95% CI -0.008 to -0.003 %-signal/month, $t(120) = -4.47$, $P_{FDR} < 0.001$). This is an intriguing phenomenon in which this waning region responds more strongly to limbs in 5–9-years-olds than in 13–17-year-olds (Fig. 3b, right, yellow). There were no statistically significant changes in responses to other categories (Fig. 3b, right, Supplementary Fig. 11 and Supplementary Table 11, full statistics). These changes in response amplitudes resulted in significant decreases in limb selectivity (Fig. 3b, left and Extended Data Fig. 2) as well as significant increases in face selectivity (Fig. 3b,
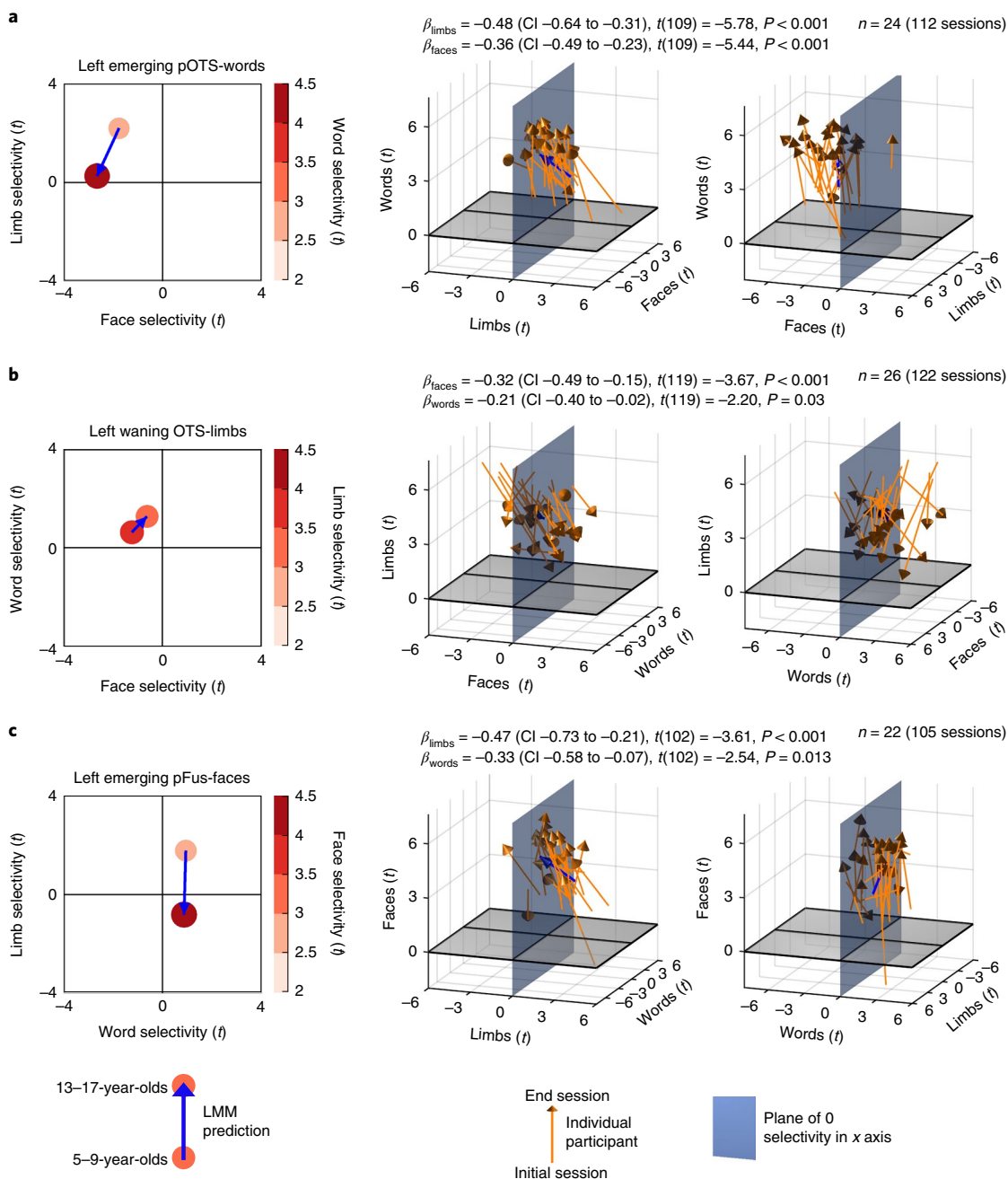
Fig. 4 | Developmental changes in word, face and limb selectivities are linked. a–c, Left: LMM prediction (circle) of category selectivity to words (**a**), limbs (**b**) and faces (**c**) versus selectivity to the other two categories in 5–9-year-olds and 13–17-year-olds. Middle: individual participant data visualized in 3D. In each panel, the variable on the z axis is related to the x and y variables, and this relationship is quantified in the model; β values, 95% CIs, t values, degrees of freedom and P values are shown at the top. Full statistics in Supplementary Table 13; Orange arrows show individual child data. Blue arrows show LMM prediction (same as left panel). Right: rotated version of the plots in the middle column to enhance visibility of positive and negative values along the other horizontal axis. Right hemisphere in Extended Data Fig. 3.

left). In right waning OTS-limbs, these increases in face selectivity were also significant when limbs and words were excluded as control categories from the contrast (Extended Data Fig. 2). There were no statistically significant changes in selectivity to other categories, including bodies (Extended Data Fig. 2). Therefore, developmental decreases in limb selectivity of waning OTS-limbs reflect decreased response amplitudes to the preferred category together with increased response amplitudes to faces.

Given the profound developmental decreases in limb selectivity in both emerging and waning ROIs, we tested whether this is a general phenomenon across lateral VTC. Analyses of lateral VTC excluding voxels that were selective in the first session to categories showing development, revealed no statistically significant decreases in limb selectivity in the remainder of lateral VTC (Supplementary Fig. 12).

Are changes in face, word and limb selectivity linked? While the theory of competition does not make predictions about limb representations, it predicts that development of face and word representations are linked as they compete for cortical territory that shows a

foveal bias[13,14]. Thus, we tested whether there is a quantitative relationship between selectivities to faces, words and limbs in the emerging and waning ROIs. Model comparison of LMMs relating the selectivity to the preferred category to selectivity of the other categories revealed that, in all developing ROIs, selectivity to the preferred category was better predicted by the selectivity to the other two categories rather than just one of them (likelihood ratio tests comparing a one-predictor LMM with a two predictor LMM, left hemisphere: all $\chi^2 \geq 4.76$, $P \leq 0.029$; Supplementary Table 12). Moreover, in all developing ROIs, selectivity to the preferred category was significantly and negatively related to selectivity to the other two categories ($\beta_{category1}$ and $\beta_{category2}$ as fixed effects, participant as random effect; Fig. 4, left hemisphere and Extended Data Fig. 3, right hemisphere; statistics in figure caption and full statistics in Supplementary Table 13). For example, in emerging pOTS-words, higher word selectivity is significantly linked with both lower face and limb selectivity (Fig. 4a and Supplementary Table 13). The negative relationship between the preferred category of an ROI and the other two categories was also observed at the voxel level (Supplementary Fig. 13).

We visualized how selectivity to words, faces and limbs changes in emerging and waning ROIs. Using the LMM, we related the selectivity to the preferred category with the selectivity to the other two categories for 5–9-year-olds and for 13–17-year-olds (Fig. 4, left and Supplementary Table 14). In emerging pOTS-words, 5–9-year-olds have positive selectivity to limbs and negative selectivity to faces (Fig. 4a). By age 13–17 years, word selectivity has increased while limb selectivity has reduced to zero and face selectivity has become even more negative (Fig. 4a, left). That is, after development, selectivity to words in pOTS-words has replaced the initial selectivity for limbs, not faces. Notably, this developmental pattern is visible in individual children from their initial age (Fig. 4a, middle and right, arrow bases) to their end age (Fig. 4a, middle and right, arrowheads). Similarly, in emerging pFus-faces, 5–9-year-olds have positive limb selectivity but no clear preference for words (Fig. 4c, left). By age 13–17 years, as face selectivity has increased, limb selectivity is lost and there is little change to word selectivity (Fig. 4c, Extended Data Fig. 3 and Supplementary Table 14).

In waning OTS-limbs, 5–9-year-olds exhibit largely negative face selectivity and mild positive word selectivity (Fig. 4b, left and Supplementary Table 14). As limb selectivity declines by age 13–17 years, both word and face selectivity increase (Fig. 4b, left). While limb selectivity consistently declined across individuals, there was more variability in individual developmental trajectories compared with the emerging ROIs (Fig. 4b, middle, right). In some children limb selectivity was replaced by word selectivity, and in others with face selectivity.

As multiple control categories were used to determine selectivity in the prior analysis, we also sought to directly measure whether there are changes in preference between the ROI-defining category versus each of the two other developing categories. For instance, in emerging pOTS-words, we directly measured preference for words versus faces and words versus limbs in each session. We plotted this preference for each child and used LMMs to estimate the significance of developmental effects as well as estimate this preference from age 5–17 years.

Results of pairwise selectivity analyses in the developing ROIs replicate results of Fig. 4, at both the individual level (Fig. 5 and Extended Data Fig. 4, thin lines) and group level (all LMM slopes significant, except for faces versus words, in left pFus-faces; statistics in Fig. 5, Extended Data Fig. 4 and Supplementary Table 15). In emerging left pOTS-words, there is a significant age-related increase of preference for both words versus limbs and words versus faces (statistics in Fig. 5 and Supplementary Table 15). It is interesting that preference for words versus limbs estimated from the LMM is negative at age 5 years but flips to be positive around age 9 years (Fig. 5a). However, there is no such flip in word versus face preference, which

steadily becomes more positive from age 5 years onwards. This suggests that, in emerging left pOTS-words, the initial preference for limbs versus words is flipped to a later preference for words versus limbs. Likewise, in left emerging pFus-faces, the negative preference for faces versus limbs (but not faces versus words) at age 5 is flipped to a positive preference later in childhood (Fig. 5c and Extended Data Fig. 4). Consistent with the prior analyses, waning OTS-limbs showed significant age-related decrease in both limbs versus faces (Fig. 5b, left) and limbs versus words (Fig. 5b, right, statistics in Fig. 5 and Supplementary Table 15). In fact, waning OTS-limbs starts with positive preferences to limbs versus faces and limbs versus words at age 5 years but loses both preferences by age 17 years. That is, at age 17 years, this waning region has no clear preference for these three categories. In sum, these analyses provide evidence that, in emerging word- and face-selective regions, earlier limb selectivity is repurposed into word and face selectivity, respectively.

Discussion

Our longitudinal measurements in children using a large range of ecologically relevant categories reveal new insights into the functional development of high-level visual cortex. We show that, while face- and word-selective regions grow and become more selective, limb-selective regions shrink during childhood development and lose their selectivity to limbs. Importantly, the decrease in limb selectivity is directly linked to the increase in word and face selectivity, providing evidence for cortical recycling during childhood development. We discuss each of these findings and their implications below.

A surprising finding from our study is that childhood development is not only associated with growth of category-selective regions and increases in selectivity, but also involves loss of selectivity. That is, in addition to finding growth and increased selectivity of face- and word-selective regions in VTC as individual children develop, consistent with cross-sectional studies[8,12,16,21–23], we find that limb-selective regions in VTC shrink during childhood development and lose their selectivity to limbs. As such, our results show that young children's VTC is actually more selective for limbs than it is later in childhood. No significant developmental changes were found for body-selective volume (Fig. 1c), consistent with previous findings[9,22].

Our results provide empirical evidence for recycling[14] of category selectivity in high-level visual cortex during childhood. However, contrary to previous predictions[13,14] that face selectivity is recycled to word selectivity during development, our results show that limb selectivity is recycled to both word and face selectivity. This recycling occurs via a mechanism of decreasing responses to limbs and increasing responses to both faces and words at the ROI and voxel levels. While we find that, in emerging left pFus-faces, not only responses to faces but also responses to words increase with age (Fig. 3c, right), future research is required to determine whether this finding reflects competition for foveal resources between words and faces[11,13]. Crucially, our results reveal no evidence that initial face selectivity is flipped into word selectivity with age but show that limb selectivity is repurposed into word and face selectivity with development. Future research is needed to determine whether this recycling also occurs at the single-neuron level[24].

Critically, our results require a rethinking of prevailing developmental theories that propose that cortical development involves sculpting of new representations upon general-purpose cortex[8,9]. First, in contrast to the prevailing view suggesting that children's VTC is indistinctive[8–10,16,25], the present data show that young children's VTC is more selective to limbs than it is later in childhood. Second, our data suggest that, during childhood, cortical selectivity can change from one category to another.

These results generate new questions for future research. First, why do young children exhibit large VTC representations for limbs?
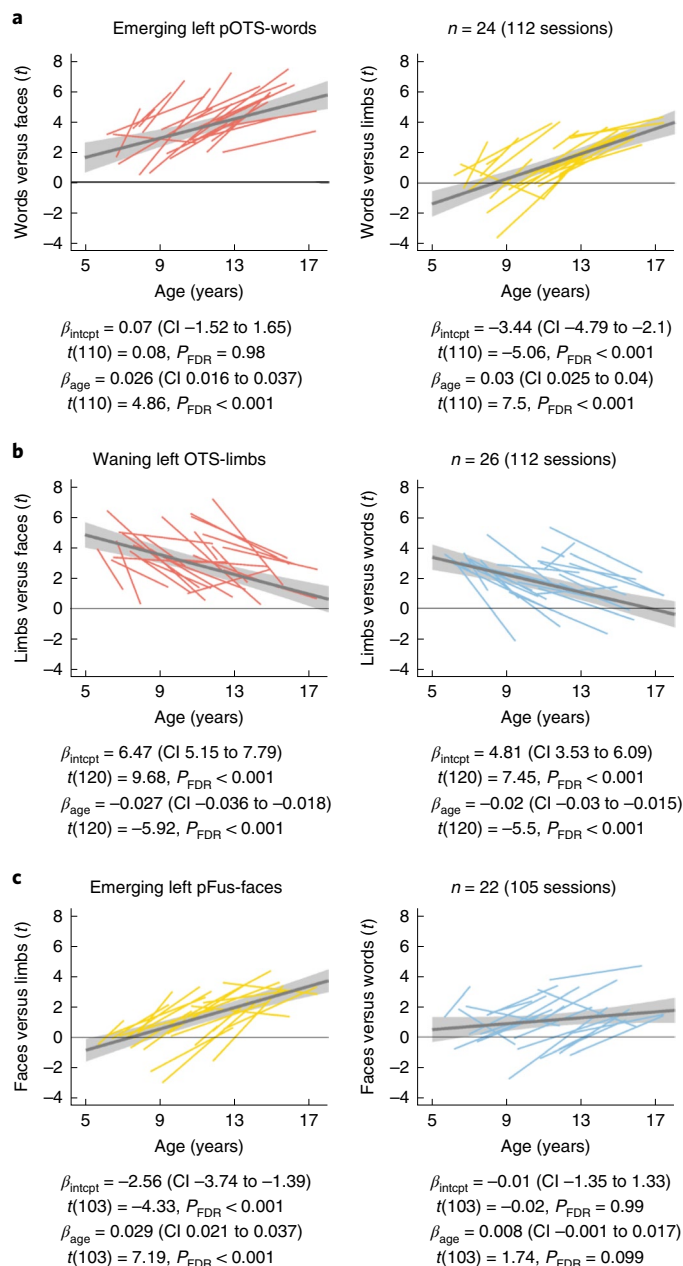
Fig. 5 | In emerging category-selective ROIs, selectivity flips from preference for limbs at age 5 years to preference for either words or faces, respectively, later in childhood. Each plot shows the pairwise preference of the ROI-defining category versus each of the other two developing categories as a function of age, with individual participant data showing the pairwise preference from the initial to end session (thin lines), LMM prediction of pairwise preference based on data from all sessions (grey line) and 95% CI (grey shading). LMM results (intercept, $\beta_{intercept}$; slope, β_{age}; rate of change in preference, t per month) and their significance are reported under each panel. **a**, Emerging left pOTS-words ($n = 24$ participants, $n = 112$ sessions). Left: words versus faces (red); right: words versus limbs (yellow). **b**, Waning left OTS-limbs ($n = 26$ participants, $n = 122$ sessions). Left: limbs versus faces (red). Right: limbs versus words (blue). **c**, Emerging pFus-faces ($n = 22$ participants, $n = 105$ sessions). Left: faces versus limbs (yellow). Right: faces versus words (blue). Right hemisphere data in Extended Data Fig. 4. Statistics in Supplementary Table 15.

Some clues to this question can be gleaned from behavioural studies that examined what infants and toddlers look at in natural settings. These studies discovered that young infants (≤6 months)

look at faces more than hands, but older infants and toddlers (1–2-year-olds) look at hands more than faces[26–28]. Developmental psychologists have hypothesized that this change in the child's 'visual diet' may be related to multiple factors including the child's mobility[28], dexterity at manipulating objects[26,28] and use of communicative information in gestures[29]. Based on this developmental literature and our findings, we speculate that, much like baby teeth changing to permanent teeth, as children's diet and size changes, cortical recycling in VTC may reflect adjustment to changing visual demands during childhood. Thus, we propose an intriguing new hypothesis that cortical recycling in VTC co-occurs with changes in the saliency and frequency of visual stimuli that are socially and communicatively relevant, moving from hands and gestures in toddlers to faces and words in school-age children and teens. Future research can test this hypothesis by determining what children and teens look at together with computational modelling to test the effect of different visual diets on emerging VTC representations[30–32].

Second, what are the behavioural implications of this cortical recycling? Prior research has found that developmental improvements in face recognition ability[8,33] and face discriminability[34] are linked with developmental increases in face selectivity[8,33] and neural sensitivity to faces[34] in VTC, respectively. Likewise, developmental improvements in reading proficiency are linked with (i) increases in neural selectivity to words[9,35,36] and (ii) increases in the informativeness of distributed responses to words in VTC[12]. An open question for future research is what might be the behavioral ramifications of the developmental decreases in responses to limbs in VTC.

The present findings have important implications for understanding typical[37] and atypical[38–41] brain development. First, these data fill a key gap in knowledge by quantifying the rate of the development of category selectivity from young children to teens. Thus, they offer a foundation for using fMRI to assess developmental and learning deficiencies, especially those related to reading[42] and social perception[38,39]. We acknowledge that a limitation of the current study is the variability in data acquisition with regard to the time interval between scans and number of scans per child. Thus, we emphasize the necessity for future longitudinal research with more regular age sampling spanning a large duration in both typical and atypical populations, which will allow further in-depth analyses of cortical recycling at finer spatial scales, such as on the level of individual voxels. Second, it will be important to determine whether there is a critical period during development in which cortical recycling can occur[10,43] and whether this period is particularly protracted in lateral VTC, which overlaps foveal representations[11,15,19]. Third, childhood visual deprivation[40] or brain lesions[41] may affect cortical recycling and the emergence of category representations in VTC in atypical populations[44–47]. Future longitudinal studies measuring cortical development in children who have vastly different visual experience with faces, hands and written words (for example, congenital blind, sighted and congenital deaf) will be important for determining how differing visual inputs and behavioural demands during childhood affect cortical recycling. Finally, an open question for future research is whether cortical recycling occurs in other brain systems in which earlier representations may be altered by schooling[48] or prolonged childhood experience. Some examples include learning new languages after one's mother tongue[49,50] or learning complex math[51] upon an earlier numerosity system[52,53].

In sum, the surprising finding from our study is that, contrary to the prevailing hypothesis that, during childhood, new cortical representations are formed upon unspecified cortex, we find evidence for cortical recycling in which limb representations are repurposed during childhood to represent faces and words. The discovery of cortical recycling is important because it not only provides a key advancement in understanding cortical development but also necessitates a rethinking of how cortical function develops during childhood.

Methods

Statement on ethical regulations. This study was approved by the Institutional Review Board of Stanford University and complies with all relevant ethical regulations. Prior to the start of the study, parents gave written consent, and children gave written assent. For their participation children received $30 per hour for scanning, as well as a small toy.

Participants. Children with normal or corrected-to-normal vision were recruited from local schools in the San Francisco Bay Area. The diversity of the participants reflects the makeup of the Bay Area population: 62.5% of children were white, 20% were Asian, 5% were Native Hawaiian, 5% were Hispanic and 7.5% were multiracial or from other racial/ethnic groups.

Prior to their first MRI session, children were trained in a scanner simulator to acclimate them to the scanner environment and to enhance the quality of MRI data. In the simulator, children practised lying still while watching a short movie and receiving live feedback on their motion. For subsequent scans, simulator training was repeated if necessary.

We collected data from 40 children (26 female, initial age 5–12 years, mean 8.66 years, s.d. 2.34 years; Supplementary Fig. 1). We selected this age range because (i) it captures the phase in which children start learning to read and (ii) it covers a broad age range spanning childhood and adolescence in which previous studies have documented VTC development[9,21].

Data from four children were excluded because they dropped out of the study after participating only once and thus did not provide longitudinal data. Data from seven children were excluded because their data did not pass inclusion criteria (see below). In the remaining 29 children, 29 functional sessions were excluded due to motion, 1 session due to a technical error during acquisition and 1 session due to aliasing artifacts during acquisition. Therefore, data from 128 functional sessions of 29 neurotypical children (18 female, 11 male) are reported in this study (Supplementary Fig. 1A,B provides an overview of the included and excluded sessions). Initial ages of the included children ranged from 5 to 12 years (mean 9.19 years, s.d. 2.13 years). No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous cross-sectional publications[33,54] and larger than those in previous longitudinal studies on VTC development[9].

Participants were scanned using functional and structural MRI for 1–5 years. When possible, children participated in one or two functional scans per year. Additionally, children participated in one structural MRI session per year. Each child participated in at least two and up to ten fMRI sessions (mean 4.41, s.d. 1.92) with the time interval between the first and last fMRI scan ranging from 10 months to 5 years (mean 45 months, s.d. 18 months; Supplementary Fig. 1C). Functional and anatomical scans were typically conducted on different days to avoid fatigue.

Magnetic resonance imaging. *Structural imaging.* Data were acquired at the Center for Cognitive Neurobiological Imaging at Stanford University on a 3-T GE Discovery MR750 scanner (GE Medical Systems) using a phase-array 32-channel head coil. Whole-brain anatomical scans were collected using quantitative MRI (qMRI[55]) with a spoiled gradient echo sequence using multiple flip angles (α = 4°, 10°, 20° and 30°), with repetition time (TR) of 14 ms and time to echo (TE) of 2.4 ms. The scan resolution was $0.8 \times 0.8 \times 1.0 \, mm^3$ (later resampled to 1 mm isotropic). For T1 calibration, we acquired spin-echo inversion recovery scans with an echo-planar imaging read-out, spectral spatial fat suppression and a slab inversion pulse. These scans were acquired using a TR of 3 s, in-plane resolution of 2 mm × 2 mm, slice thickness of 4 mm, 2× acceleration and time to echo of minimum full.

Functional imaging. Functional data were collected using the same scanner and head coil as the structural images. Slices were oriented parallel to the parieto-occipital sulcus. We used a simultaneous multi-slice, one-shot T2* sensitive gradient echo EPI sequence with a multiplexing factor of 3 to acquire near-whole-brain coverage (48 slices) with a field of view of 192 mm, TR of 1 s, TE of 30 ms and flip angle of 76°. Resolution was 2.4 mm isotropic.

Ten-category experiment. Participants completed three runs of a ten-category experiment[11,12,54]. During each run, participants viewed images from five domains, each comprising images from two categories: faces (adult faces, child faces), body parts (headless bodies, limbs), objects (cars, string instruments), places (corridors, houses) and characters (pseudowords, numbers). Following prior work on visual representations, we define a visual category as a set of exemplars sharing the same parts and configuration, for example, limbs[56–59], and domain as a grouping of one or more categories that share semantic association (whether or not they share visual features; for example, houses and corridors are both places but are visually dissimilar) and are thought to require distinct cortical processing mechanisms[60]. Examples of stimuli are shown in Supplementary Fig. 1D.

Images were presented in 4 s blocks at a rate of 2 Hz and did not repeat across the course of the experiment. Image blocks were intermixed with grey luminance screen baseline blocks. Blocks were counterbalanced across categories and baseline. Stimuli were greyscale and contained a phase-scrambled background generated from randomly selected images. Participants were instructed to view the images

while fixating on a central dot and to perform an oddball task by pressing a button whenever an image comprising only the phase-scrambled background appeared. Due to occasional button box malfunction, behavioural responses in the oddball task were recorded in 98 out of 128 sessions used in the analyses.

Data analysis. Data analysis was performed in MATLAB version 2017b (MathWorks, Inc.) and using the mrVista software package (https://github.com/vistalab/vistasoft/wiki/mrVista).

Inclusion criteria. In each functional session, children participated in three runs of the ten-category experiment. Criteria for inclusion of data were (i) at least two runs per session with within-run motion <2 voxels and between-run motion <3 voxels, and (ii) the child participated in at least two fMRI sessions that were at least 6 months apart.

Because only two of the three runs survived motion quality thresholds for several fMRI sessions, analyses include two runs per child per session to ensure equal amounts of data across participants and sessions. For sessions with all three runs passing motion quality criteria, two runs with lowest within-run motion were included.

Structural MRI data analysis and individual template. Quantitative whole-brain images for each child and timepoint were processed with the mrQ pipeline (https://github.com/mezera/mrQ, ref. [55]) to generate synthetic T1 brain volumes. For each child, the synthetic T1 brain volumes from their multiple timepoints were used to generate a within-participant brain volume template. Each participant's brain anatomical template was generated using the FreeSurfer Longitudinal pipeline implemented in FreeSurfer version 6.0 (https://surfer.nmr.mgh.harvard.edu/fswiki/LongitudinalProcessing, ref. [61]) The grey–white matter segmentation of each individual's within-participant brain template was manually edited to fix segmentation errors (for example, holes and handles). From this segmentation we generated an accurate cortical surface reconstruction of each participant's brain. The motivations for aligning the functional data to the within-participant-template were to: (i) enable comparison of ROIs from different timepoints in the same brain volume for each participant, and (ii) minimize potential biases which can occur from aligning longitudinal data to the anatomical volume from a single timepoint[61]. On average 2.48 (s.d. 0.69) synthetic T1s were used to generate the within-participant template (minimum 2, maximum 5). Functional data from all sessions of a participant were aligned to their within-participant brain template. In 17 participants, the last fMRI session that was included was conducted after the within-participant template had been created. These functional sessions were acquired on average 11 ± 2 months after acquisition of the last synthetic T1 that was included in the within-participant template (excluding two participants whose last T1 could not be used because of technical error during acquisition and participant motion).

Definition of anatomical VTC ROIs. On the inflated surface of each hemisphere in each participant, we defined anatomical ROIs of the lateral and medial VTC (Fig. 1b) as in previous publications[12]. We first defined the VTC anatomically and then separated it into lateral and medial VTC. The posterior border of VTC was the posterior transverse collateral sulcus (ptCoS), and the anterior border was aligned to the posterior end of the hippocampus, which typically aligns with the anterior tip of the mid-fusiform sulcus (MFS). The lateral border of VTC was the inferior temporal gyrus (ITG), and the medial border of VTC was the medial border of the collateral sulcus (CoS). Finally, VTC was divided into its lateral and medial partitions along the MFS (example in Fig. 1b).

fMRI data analysis. Functional data from each session were aligned to the individual within-participant template. Motion correction was performed both within and across functional runs. No spatial smoothing or slice-timing correction was performed. Time courses were transformed into percentage signal change by dividing each timepoint of each voxel's data by the average response across the entire run. To estimate the contribution of each of the ten conditions (corresponding to the ten image categories), a general linear model (GLM) was fit to each voxel by convolving the stimulus presentation design with the haemodynamic response function (as implemented in SPM, www.fil.ion.ucl.ac.uk/spm).

Definition of selectivity in lateral and medial VTC. We used a data-driven approach to examine the development of category selectivity in VTC. The motivation for this type of analysis was to (i) use an automated, observer-independent approach and (ii) use an approach that does not require clustered activations. In each participant, we assessed selectivity to each category in anatomically defined lateral and medial VTC ROIs (Fig. 1b). Selectivity was defined as $t > 3$ (voxel level) for the contrast of interest. We also performed a complementary threshold-independent analysis in constant-sized regions (see control analyses below). Category contrasts (Fig. 1) were computed by contrasting responses to each category with all other categories from all other domains (that is, words versus all other categories except numbers). We used this approach for defining contrasts to ensure that all categories are contrasted to the same number of control categories and contrasts are not biased

to any particular category. We also computed contrasts for domains (faces, body parts, characters, objects and places) (see supplemental analyses in Supplementary Fig. 4). For domain contrasts, responses to both categories of each domain were contrasted with stimuli from all the other domains (that is, characters versus all others).

Control analysis on categories included in the contrasts. Results of the analysis related to Fig. 1 revealed a bilateral decrease in limb selectivity in addition to increases in face and word selectivity with age. Given that face-, limb- and word-selective regions neighbour in the brain, it is possible that the observed decrease in limb selectivity may partially be driven by the inclusion of categories showing age-related increases in selectivity (for example, faces and words) as control categories in the contrast (see above). Therefore, to ensure that the decrease in limb selectivity and the increase in face and word selectivity are not driven by including the respective other two categories as controls in the contrasts, we repeated all analyses related to Figs. 1 and 3 but excluding these three categories from being control categories in all contrasts. That is, we contrasted each category versus all other categories excluding limbs, words, adult faces, child faces and the other category from the same domain. For instance, responses to limbs were contrasted with responses to cars, string instruments, houses, corridors and numbers. Results are shown in Extended Data Figs. 1 and 2 and Supplementary Fig. 6.

Definition of functional ROIs. To examine the anatomical specificity of the observed development of selectivity (Fig. 1), we defined word-, limb- and face-selective functional regions in each participant (Fig. 2a–c). For the definition of functional ROIs, a threshold of $t > 3$ (voxel level) was used. All ROIs were defined on each participant's native cortical surface generated from the within-participant brain template. Word-selective regions were defined as above-threshold clusters for the contrast words versus all categories except numbers that straddled the OTS. The more anterior cluster was defined as mOTS-words, and the posterior cluster as pOTS-words. These clusters are also referred to as the visual word form area[4] (VWFA 1 and 2, respectively). The limb-selective region was defined as above-threshold cluster of voxels for the contrast limbs versus all categories except bodies straddling the OTS and was labelled OTS-limbs. This cluster is also referred to as the fusiform body area[2] (FBA). Face-selective regions were defined using the contrast faces (adult and child) versus all other categories, because we observed similar development for both face types in the analysis related to Fig. 1. Face-selective clusters were defined as above-threshold clusters on the lateral fusiform gyrus. The more anterior cluster typically aligns with the anterior tip of the MFS and was defined as mFus-faces[19], while the more posterior cluster was defined as pFus-faces[19]. These two face-selective clusters are also referred to as the fusiform face area[1] (FFA 1 and 2, respectively).

For supplemental analyses (Supplementary Figs. 4 and 9) we also defined place-, character- and combined body-part-selective regions. Place-selective regions were defined as above-threshold clusters on the collateral sulcus that respond to houses and corridors versus all other categories. This region is also referred to as parahippocampal place area[3] (PPA). Character-selective regions were defined as above-threshold clusters on the OTS that responded more strongly to words and numbers versus all other categories. Similarly, a body-part-selective region was defined as above-threshold clusters on the OTS that responded more strongly to bodies and limbs than the other categories.

Emerging and waning ROIs. To characterize the selectivity and responses within VTC regions that changed with development, we defined ROIs in each participant's VTC that: (i) either gained selectivity to faces or words or (ii) lost selectivity to limbs during childhood. To do so, we defined emerging face and word ROIs as well as waning limb ROIs. For words and faces, where selective voxels emerge during development, the emerging ROI is the region that was selective to faces (or words) at the end timepoint but was not selective to that category at the initial timepoint (Fig. 2a,c). We call the region that is the difference between the end and initial ROIs emerging pOTS-words and emerging pFus-faces, respectively. For limbs, where selectivity declines during development, we call the difference between the initial and end ROIs the waning OTS-limbs. That is, waning OTS-limbs is the region that was within the ROI at the initial timepoint but not at the end timepoint (Fig. 2b). For a given participant, the initial ROI corresponds to the first fMRI session in which the ROI could be identified while the end ROI corresponds to the last fMRI session in which the ROI could be identified.

Control analysis in independent ring-shaped ROIs. The emerging and waning ROIs were defined from individual participant's functional ROIs on their native brain anatomy and thus capture precisely the part of cortex that undergoes development. As these ROIs are not completely independent from the original ROIs, we sought to validate the results in an independent manner. Thus, we conducted a complementary analysis to evaluate responses in independent ring ROIs. First, we created two disk ROIs centred at the centre coordinate of the initial functional ROI. One ROI was sized to match the surface area of the initial ROI and the other to match the corresponding end ROI. Then, we defined the area in between these two disk ROIs as the independent ring-shaped ROI.

Within each ring-shaped ROI, we measured responses to the ten categories as well as selectivity to each category (Supplementary Fig. 10). Importantly, this analysis replicated the significant decrease in limb selectivity in waning limb-ROIs, emerging left pOTS-words and emerging left pFus-faces. In right pFus-faces, we found a similar trend for decreasing limb selectivity (see legend of Supplementary Fig. 10 for statistics). In left pOTS-words we find a similar trend for increasing word selectivity (Supplementary Fig. 10A), and in right pFus-faces we find a similar trend for increasing face selectivity (Supplementary Fig. 10D,E). We note that a limitation of this approach is that, while the ring analysis guarantees independence, it does not capture exactly the developing tissue, as the actual developing ROIs are not ring shaped (Fig. 2a–c, see example ROIs). Therefore, this approach is less accurate for assessing developmental changes in VTC.

Control analysis in a constant number of lateral VTC voxels. In the analyses in Fig. 1, we evaluated category selectivity by estimating the number of voxels within an anatomical lateral VTC ROI that significantly responded to one category (versus all other categories except the other category from the same domain, threshold $t > 3$, voxel level). While this threshold guarantees significant selectivity, ultimately the threshold value is a number decided by the experimenter.

To ensure that findings of developmental effects do not depend on the threshold used, we performed a complementary analysis of category selectivity in lateral VTC across development which did not depend on a threshold. Across development, we kept the number of voxels constant and evaluated their mean selectivity (t value) to a category. Specifically, at each timepoint, we selected the 20% most selective voxels (that is, the voxels with the highest t values) for each category contrast in lateral VTC and calculated their mean t value (Supplementary Fig. 5). The 20% most selective voxels are determined in each session independently, to avoid biasing the selection of voxels to a specific timepoint. Then, we used LMMs to determine whether the mean selectivity of these voxels changes over time. LMMs can be expressed as: t value ~ age in months + (1|participant).

Results of this analysis largely replicate the main finding presented in Fig. 1c, and reveal: (i) a significant increase in word selectivity in the left hemisphere (see legend of Supplementary Fig. 5 for statistics), (ii) a bilateral decrease in limb selectivity, (iii) an increase in selectivity to adult faces in the right hemisphere, (iv) a bilateral increase in selectivity to child faces and (v) an increase in selectivity to houses in the left hemisphere.

Control analysis on response reliability in V1. To assess the reliability of responses, we conducted a multivariate pattern analysis (MVPA[62]) in V1 ROIs defined by the Glasser atlas[63]. For each voxel, the response amplitude (β value) for each category was estimated from the GLM of each run. These vectors of responses, also called multivoxel patterns (MVPs), were generated independently for each of the two runs in each session. We then calculated correlations between pairs of MVPs (run 1 to run 2) to each category combination, resulting in 10×10 representational similarity matrices. We calculated the mean reliability across all categories (mean of the on-diagonal correlations) as a measure for the reliability of responses, as it provides a measurement for the similarity of responses to the same category from one functional run to another. Next, we tested whether there is a relationship between the reliability of responses and age. To this end, we ran LMMs predicting the response reliability by age using random intercept models with the grouping variable participant (response reliability ~ age + (1|participant)). Results of this analysis are shown in Supplementary Fig. 7.

Is the decrease in limb selectivity uniform in lateral VTC? As we found accumulating evidence for reductions in limb selectivity in lateral VTC, an open question is whether the decrease in limb selectivity occurs across the entire lateral VTC or whether it is restricted to the regions in which selectivity changes across development. To this end, we aimed to assess development of limb selectivity in VTC while excluding voxels that were selective to categories showing significant development (Fig. 1). Thus, we identified in each participant's first session the voxels selective for these categories, excluded them from lateral VTC, and then measured limb selectivity across development in the remaining lateral VTC voxels. Next, we used LMMs to test whether there was a significant decrease in the mean selectivity to limbs in these remaining lateral VTC voxels (Supplementary Fig. 12). While there was a trend for a decrease in selectivity in both hemispheres, the effects were not significant after FDR correction (left: slope = −0.0039 t *per* month, $P_{FDR} = 0.16$; right: slope = −0.0036 t per month, $P_{FDR} = 0.16$, FDR corrected).

Can cortical recycling be measured at the voxel level? To assess whether cortical recycling is also observable at the voxel level, we measured the selectivity for limbs, faces and words in each session for each voxel of the waning and emerging parts of category-selective ROIs. Using this data, we examined the relationship between the voxel selectivity to the category that defines the ROI as a function of voxel selectivity to the two other categories (that is, in emerging pOTS-words we related word selectivity to limb and face selectivity: word selectivity ~ limb selectivity + face selectivity +(1|session)). To quantify the relationship, we ran a separate LMM for each participant and ROI. Thus, the number of independent measurements in each LMM is the number of voxels in the emerging/waning ROI, and the grouping

variable in the LMM is 'session'. We then summarized the individual participant slopes from the LMM for each ROI in a boxplot (Supplementary Fig. 13)

Statistical analyses. LMMs were used for statistical analyses because: (i) the data have a hierarchical structure with sessions being nested within each participant, and (ii) sessions were unevenly distributed across time (Supplementary Fig. 1A). Models were fitted using the 'fitlme' function in MATLAB version 2017b (MathWorks, Inc.). In initial analyses, the fit of random-intercept models, which allow intercepts to vary across participants, were compared with the fit of random-slope models, which allow both intercepts and slopes to vary across participants. Results revealed that a random-intercept model fitted the data best in the majority of cases. Thus, to enable comparability across analyses, LMMs with random intercepts were used throughout the analyses presented herein. Due to the relatively large number of samples, data distribution was assumed to be normal, but this was not formally tested.

LMMs related to Fig. 1 can be expressed as: volume of selective activation for a category in mm³ ~ age in months + (1|participant), in which volume of selective activation is the response variable, age in months is a continuous predictor (fixed effect) and the term (1|participant) indicates that random intercepts are used for the grouping variable participant. The slopes of the LMMs are plotted in Fig. 1c. Statistics were run on the complete dataset including 128 sessions. Boxplots in Fig. 1d showing volume of selective activation in different age groups based on a subset of the data are used for visualization purposes only and not to evaluate statistical significance. One session per child per age group is included in the boxplot. For 5–9-year-olds, each participant's first session was included, and for 13–17-year-olds, each participant's last session was included. To confirm the validity of the choice of these age groups for the boxplots in Fig. 1d (5–9-year-olds and 13–17-year-olds), we used the LMMs (Fig. 1c) to predict the size of category-selective activation for the mean age in years of the participants in each of the two age groups. This predicted size is indicated as a black diamond in the boxplots. Estimated mean size (diamonds) from the LMM corresponded well with the medians of the boxplots, thus validating the grouping of the participants in the boxplots.

To estimate changes in the size of category- and domain-selective ROIs (Fig. 2d and Supplementary Fig. 9) we used LMMs specified as: ROI size in mm³ ~ age in months + (1|participant). Similarly, to estimate changes in the surface area of category-selective ROIs (Supplementary Fig. 8), we used LMMs specified as: ROI surface area in mm² ~ age in months + (1|participant).

To evaluate significance of the developmental changes in emerging and waning ROIs, we used two separate LMMs: (i) selectivity changes across development (Fig. 3-left) were modelled as *t* value ~ age in months + (1|participant), and (ii) changes in responses across development (Fig. 3-right) as % signal (β) ~ age in months + (1|participant). Percent signal refers to the change in response for a certain condition relative to a blank baseline. Selectivity refers to contrasting responses to different categories (see above definition of contrasts and thresholds). Selectivity and response values were obtained for each voxel, and LMMs were fitted on the average value in an ROI. Boxplots in Fig. 3 show mean responses for each of the ten categories in 5–9-year-olds and 13–17-year-olds. One session per child per age group is included in the boxplot. For 5–9-year-old, each participant's first session was included, and for 13–17-year-olds, each participant's last session was included. Boxplots are used for visualizing response amplitudes (units of %-signal) in different age groups but not to evaluate statistical significance, which was done using the LMM. The selection of age groups in the boxplots is validated by plotting the LMM's prediction of the parameter of interest for the mean age of participants in each age group (compare black diamonds with measured median).

LMM analyses related to Fig. 4 tested whether changes in limb, face and word selectivity in the developing ROIs were significantly related to each other. For each emerging and waning ROI, we tested whether selectivity for one category (for example, word selectivity in emerging pOTS-words) was predicted by selectivity to the other categories (for example, limb and face selectivity in emerging pOTS-words). We also tested whether an LMM with two predictors (for example, predicting word selectivity from both limb and face selectivity) is a better model than an LMM with one predictor (for example, predicting word selectivity just from face selectivity) using a likelihood ratio test. If the likelihood ratio test confirmed that both predictors contributed significantly to the model fit, both predictors were included in the analysis. Parameters of LMMs and likelihood ratio tests are reported in Supplementary Tables 12–14. Tables are grouped by analysis type.

We also performed an analysis in which we directly evaluated the pairwise contrasts between words, faces and limbs in the developing ROIs for each child and session. For instance, in emerging pOTS-words, we directly contrasted (i) words versus faces, and (ii) words versus limbs. We used LMMs to test whether and how a region's preference for one category over the other changes with age. For each emerging and waning ROI and each contrast we ran a LMM of the format: *t* value ~ age in months + (1|participant). Left hemisphere data are presented in Fig. 5, right hemisphere data are presented in Extended Data Fig. 4 and full statistics are in Supplementary Table 15.

The reported statistical tests are two tailed. FDR correction following the procedure of Benjamini and Hochberg[20] as implemented in MATLAB version

2017b (MathWorks, Inc.) was applied to correct the statistical significance taking into account multiple comparisons related to each analysis.

Data availability
The data required to generate the main figures are available in the GitHub repository (https://github.com/VPNL/corticalRecycling). Due to the large size of the raw data, it will be made available from the corresponding author upon request.

Code availability
Code is available at https://github.com/VPNL/corticalRecycling.

References
1. Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**, 4302–4311 (1997).
2. Peelen, M. V. & Downing, P. E. Selectivity for the human body in the fusiform gyrus. *J. Neurophysiol.* **93**, 603–608 (2005).
3. Epstein, R. & Kanwisher, N. A cortical representation of the local visual environment. *Nature* **392**, 598–601 (1998).
4. Cohen, L. et al. The visual word form area. Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* **123**, 291–307 (2000).
5. Deen, B. et al. Organization of high-level visual cortex in human infants. *Nat. Commun.* **8**, 13995 (2017).
6. de Heering, A. & Rossion, B. Rapid categorization of natural face images in the infant right hemisphere. *eLife* **4**, e06564 (2015).
7. Livingstone, M. S. et al. Development of the macaque face-patch system. *Nat. Commun.* **8**, 14897 (2017).
8. Golarai, G. et al. Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nat. Neurosci.* **10**, 512–522 (2007).
9. Dehaene-Lambertz, G., Monzalvo, K. & Dehaene, S. The emergence of the visual word form: longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLoS Biol.* **16**, 1–34 (2018).
10. Srihasam, K., Vincent, J. L. & Livingstone, M. S. Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nat. Neurosci.* **17**, 1776–1783 (2014).
11. Gomez, J., Natu, V., Jeska, B., Barnett, M. & Grill-Spector, K. Development differentially sculpts receptive fields across early and high-level human visual cortex. *Nat. Commun.* **9**, 788 (2018).
12. Nordt, M. et al. Learning to read increases the informativeness of distributed ventral temporal responses. *Cereb. Cortex* https://doi.org/10.1093/cercor/bhy178 (2019).
13. Behrmann, M. & Plaut, D. C. A vision of graded hemispheric specialization. *Ann. N. Y. Acad. Sci.* **1359**, 30–46 (2015).
14. Dehaene, S., Cohen, L., Morais, J. & Kolinsky, R. Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* **16**, 234–244 (2015).
15. Levy, I., Hasson, U., Avidan, G., Hendler, T. & Malach, R. Center–periphery organization of human object areas. *Nat. Neurosci.* **4**, 533–539 (2001).
16. Cantlon, J. F., Pinel, P., Dehaene, S. & Pelphrey, K. A. Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cereb. Cortex* **21**, 191–199 (2011).
17. Dehaene, S. et al. How learning to read changes cortical networks for vision and language. *Science* **1359**, 1359–1364 (2010).
18. Frost, M. A. & Goebel, R. Measuring structural-functional correspondence: spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage* **59**, 1369–1381 (2012).
19. Weiner, K. S. et al. The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *Neuroimage* **84**, 453–465 (2014).
20. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–3300 (1995).
21. Scherf, K. S., Behrmann, M., Humphreys, K. & Luna, B. Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Dev. Sci.* **10**, F15–30 (2007).
22. Peelen, M. V., Glaser, B., Vuilleumier, P. & Eliez, S. Differential development of selectivity for faces and bodies in the fusiform gyrus. *Dev. Sci.* **12**, 16–25 (2009).

23. Golarai, G., Liberman, A., Yoon, J. M. & Grill-Spector, K. Differential development of the ventral visual cortex extends through adolescence. *Front. Hum. Neurosci.* **3**, 80 (2010).

24. Kobatake, E., Wang, G. & Tanaka, K. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J. Neurophysiol.* **80**, 324–330 (1998).

25. Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R. & Livingstone, M. S. Seeing faces is necessary for face-domain formation. *Nat. Neurosci.* **20**, 1404–1412 (2017).

26. Fausey, C. M., Jayaraman, S. & Smith, L. B. From faces to hands: changing visual input in the first two years. *Cognition* **152**, 101–107 (2016).

27. Frank, M. C., Vul, E. & Saxe, R. Measuring the development of social attention using free-viewing. *Infancy* **17**, 355–375 (2012).

28. Long, B., Kachergis, G., Agrawal, K., & Frank, M. C. Detecting social information in a dense database of infants' natural visual experience. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/z7tdg (2020).

29. Liszkowski, U., Carpenter, M. & Tomasello, M. Pointing out new news, old news, and absent referents at 12 months of age. *Dev. Sci.* **10**, F1–F7 (2007).

30. Haber, N., Mrowca, D., Wang, S., Fei-Fei, L. & Yamins, D. L. K. Learning to play with intrinsically-motivated, self-aware agents. *Adv. Neural Inf. Process. Syst.* **31**, 8388–8399 (2018).

31. Khaligh-Razavi, S. M. & Kriegeskorte, N. Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Comput. Biol.* **10**, e1003915 (2014).

32. Zhuang, C.; Zhai, A. L.; Yamins, D. in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2019).

33. Gomez, J. et al. Microstructural proliferation in human cortex is coupled with the development of face processing. *Science* **355**, 68–71 (2017).

34. Natu, V. S. et al. Development of neural sensitivity to face identity correlates with perceptual discriminability. *J. Neurosci.* **36**, 10893–10907 (2016).

35. Wandell, B. A., Rauschecker, A. M. & Yeatman, J. D. Learning to see words. *Annu. Rev. Psychol.* **63**, 31–53 (2012).

36. Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K. & Wandell, B. A. The development of cortical sensitivity to visual word forms. *J. Cogn. Neurosci.* **23**, 2387–2399 (2011).

37. Feldstein Ewing, S. W., Bjork, J. M. & Luciana, M. Implications of the ABCD study for developmental neuroscience. *Dev. Cogn. Neurosci.* **32**, 161–164 (2018).

38. Constantino, J. N. et al. Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature* **547**, 340–344 (2017).

39. Duchaine, B. C. & Nakayama, K. Developmental prosopagnosia: a window to content-specific face processing. *Curr. Opin. Neurobiol.* **16**, 166–173 (2006).

40. Amedi, A., Raz, N., Pianka, P., Malach, R. & Zohary, E. Early 'visual' cortex activation correlates with superior verbal memory performance in the blind. *Nat. Neurosci.* **6**, 758–766 (2003).

41. Liu, T. T. et al. Successful reorganization of category-selective visual cortex following occipito-temporal lobectomy in childhood. *Cell Rep.* **24**, 1113–1122. e6 (2018).

42. Norton, E. S., Beach, S. D. & Gabrieli, J. D. E. Neurobiology of dyslexia. *Curr. Opin. Neurobiol.* **30**, 73–78 (2015).

43. Srihasam, K., Mandeville, J. B., Morocz, I. A., Sullivan, K. J. & Livingstone, M. S. Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron* **73**, 608–619 (2012).

44. Büchel, C., Price, C. & Friston, K. A multimodal language region in the ventral visual pathway. *Nature* **394**, 274–277 (1998).

45. Emmorey, K., McCullough, S. & Weisberg, J. Neural correlates of fingerspelling, text, and sign processing in deaf American sign language–English bilinguals. *Lang. Cogn. Neurosci.* **30**, 749–767 (2015).

46. Bi, Y., Wang, X. & Caramazza, A. Object domain and modality in the ventral visual pathway. *Trends Cogn. Sci.* **20**, 282–290 (2016).

47. Reich, L., Szwed, M., Cohen, L. & Amedi, A. A ventral visual stream reading center independent of visual experience. *Curr. Biol.* **21**, 363–368 (2011).

48. Dehaene, S. & Cohen, L. Cultural recycling of cortical maps. *Neuron* **56**, 384–398 (2007).

49. Lucas, T. H., McKhann, G. M. & Ojemann, G. A. Functional separation of languages in the bilingual brain: a comparison of electrical stimulation language mapping in 25 bilingual patients and 117 monolingual control patients. *J. Neurosurg.* **101**, 449–457 (2004).

50. Green, D. W., Crinion, J. & Price, C. J. Convergence, degeneracy, and control. *Lang. Learn.* **56**, 99–125 (2006).

51. Amalric, M. & Dehaene, S. Origins of the brain networks for advanced mathematics in expert mathematicians. *Proc. Natl Acad. Sci. USA* **113**, 4909–4917 (2016).

52. Kersey, A. J. & Cantlon, J. F. Neural tuning to numerosity relates to perceptual tuning in 3–6-year-old children. *J. Neurosci.* **37**, 512–522 (2017).

53. Pica, P., Lemer, C., Izard, V. & Dehaene, S. Exact and approximate arithmetic in an Amazonian indigene group. *Science* **306**, 499–503 (2004).

54. Natu, V. S. et al. Apparent thinning of human visual cortex during childhood is associated with myelination. *Proc. Natl Acad. Sci. USA* **116**, 20750–20759 (2019).

55. Mezer, A. et al. Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nat. Med.* **19**, 1667–1672 (2013).

56. Grill-Spector, K. & Kanwisher, N. Visual recognition: as soon as you know it is there, you know what it is. *Psychol. Sci.* **16**, 152–160 (2005).

57. Weiner, K. S. & Grill-Spector, K. Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage* **52**, 1559–1573 (2010).

58. Weiner, K. S., Sayres, R., Vinberg, J. & Grill-Spector, K. fMRI-adaptation and category selectivity in human ventral temporal cortex: regional differences across time scales. *J. Neurophysiol.* **103**, 3349–3365 (2010).

59. Stigliani, A., Weiner, K. S. & Grill-Spector, K. Temporal processing capacity in high-level visual cortex is domain specific. *J. Neurosci.* **35**, 12412–12424 (2015).

60. Kanwisher, N. Domain specificity in face perception. *Nat. Neurosci.* **3**, 759–763 (2000).

61. Reuter, M., Schmansky, N. J., Rosas, H. D. & Fischl, B. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* **61**, 1402–1418 (2012).

62. Haxby, J. V. et al. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293**, 2425–2430 (2001).

63. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).

Acknowledgements

Author contributions

M.N. collected data, developed and coded the analysis pipeline, analysed the data and wrote the manuscript. V.S.N. and J.G. designed the experiment, collected data and contributed to the manuscript. A.A.R. collected the data, contributed to data analysis and contributed to the manuscript. D.F. and H.K. collected the data and contributed to the manuscript. K.G.-S. designed the experiment, contributed to the analysis pipeline and data analyses and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41562-021-01141-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-021-01141-5.

Correspondence and requests for materials should be addressed to K.G.-S.

Peer review information *Nature Human Behaviour* thanks Marius Peelen and Frank Tong for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
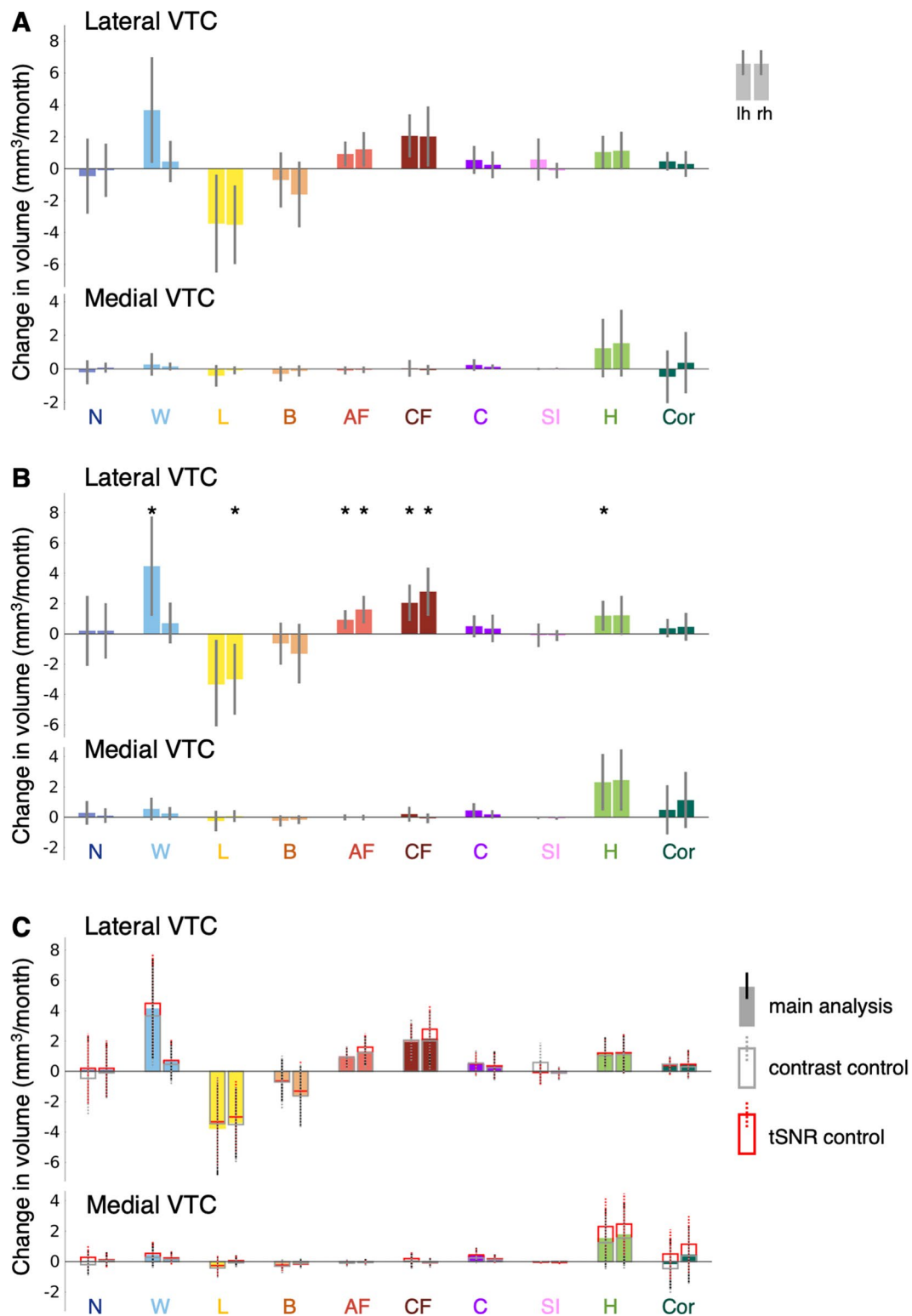
Extended Data Fig. 1 | Control analyses examining developmental increases and decreases in category-selective activation in lateral VTC. LMM slopes indicating change in category-selective activation volume per month (n = 128 sessions, 29 children). Error bars: 95% CI. If the CI does not cross the y = 0 line, this indicates that the slope is significantly different than 0 (before FDR-correction). (**a**) Slopes for the age predictor for models in which adult faces, child faces, limbs, and words are excluded as control categories from the contrast. No effects survive FDR-correction. (**b**) Slopes for the age predictor for models including both age and time series signal-to-noise ratio (tSNR) as predictors. Significant development after FDR-correction (p < 0.05) is indicated by asterisks. (**c**) Slopes for the main analysis (filled bars), the contrast control (open bars with gray outline) and the tSNR control (open bars with red outline) are overlaid to illustrate the changes in effect size across the different analyses. Full statistics in Supplementary Tables 3-4,6-7. Related to Fig. 1.
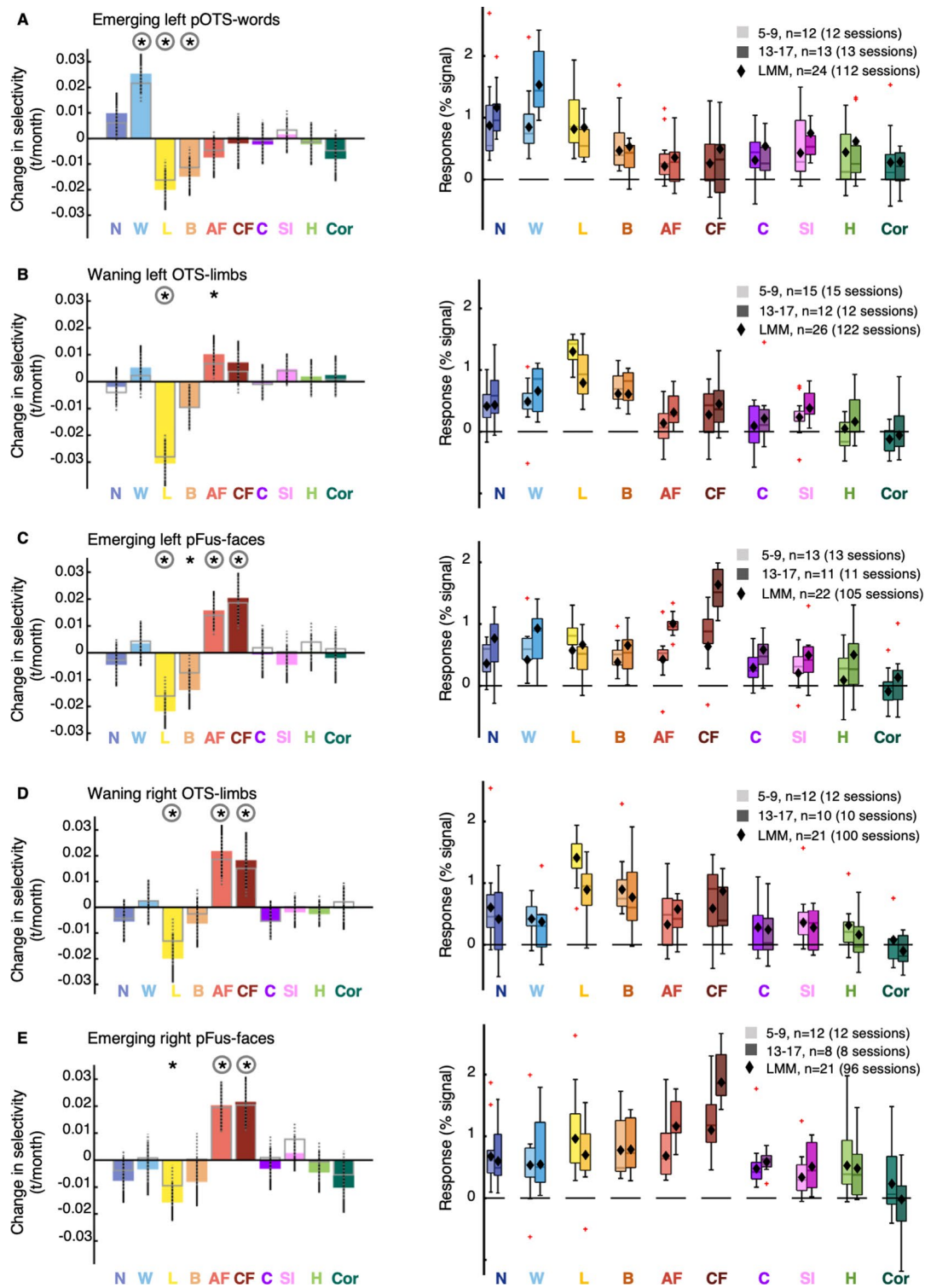
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Control analyses examining functional changes underlying the development of category-selective ROIs. *Left panel: Colored bars:* Slopes of LMMs indicating changes in selectivity by age for all 10 categories in emerging and waning ROIs. *Open bars with gray outline:* LMM slopes for contrasts in which adult faces, child faces, limbs, and words are excluded as control categories in contrasts. Error bars: 95% CI. If the CI does not cross the y = 0 line, this indicates that the slope is significantly different than 0 (before FDR-correction). *Asterisks*: significant development after FDR-correction (p < 0.05) for colored bars, *circles*: significant after FDR-correction for open bars. *Right panel:* Response amplitudes for 5–9-year-olds and 13–17-year-olds. Lighter colors indicate younger ages. One functional session per child is included per boxplot. Boxplots show the 75% and 25% percentiles (colored areas) and median (horizontal lines). Whiskers extend to the most extreme data points not considered outliers (minimum, maximum). Crosses: outliers (values more than 1.5 times the interquartile range away from the bottom or top of the box). *Black diamonds*: LMM prediction for the response at the mean age of each age group. (**a**) Left emerging pOTS-words. Left panel: n = 24 (112 sessions). (**b**) Left waning OTS-limbs. n = 26, 122 sessions. (**c**) Left emerging pFus-faces. n = 22, 105 sessions. (**d**) Right waning OTS-limbs. n = 21, 100 sessions. (**e**) Right emerging pFus-faces. n = 21, 96 sessions. Full statistics in Supplementary Tables 9-10. As we observed a significant decrease in limb-selectivity in emerging parts of word- and face-selective regions and word-, face- and limb-selective regions neighbor, we tested if emerging parts of word- and face-selective regions overlap with waning parts of the limb-selective regions. However, the overlap between the developing parts of the ROIs (difference between initial and end ROIs) assessed by the dice coefficient (DC) was small (overlap between developing parts of OTS-limbs and pFus-faces, left: DC = 0.025 ± 0.01 (mean ± SD), n = 22; right: DC = 0.026 ± 0.01, n = 18; overlap between developing parts of left OTS-limbs and pOTS-words: DC = 0.006±SD = 0.004, n = 21). Related to Fig. 3.

A Waning right OTS-limbs

B_{faces}= -0.32 (CI:-0.45,-0.19), t(97)=-4.80, p<0.001
β_{words}= -0.24 (CI:-0.46,-0.03), t(97)=-2.28, p=0.025

n=21 (100 sessions)
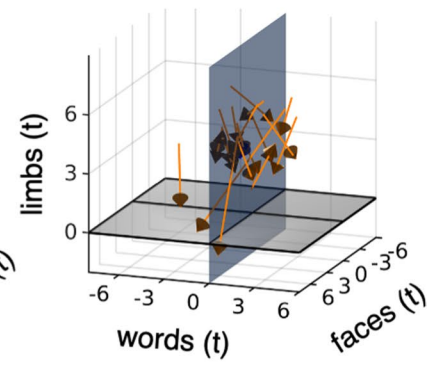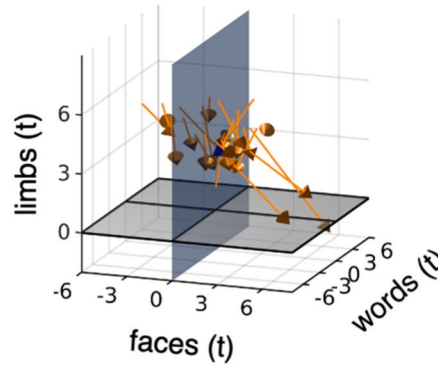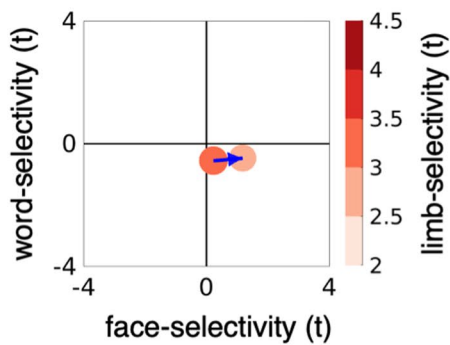
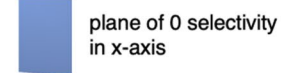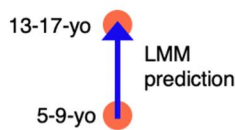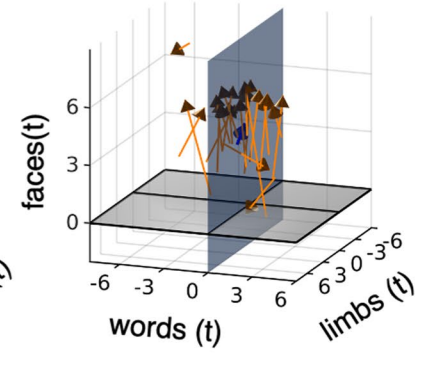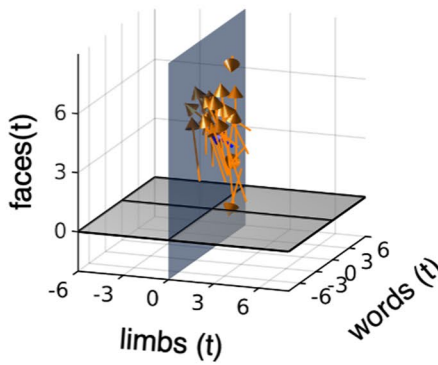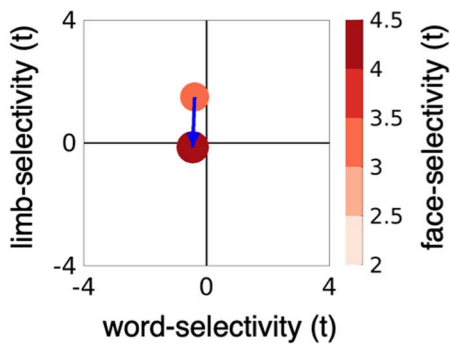## B Emerging right pFus-faces

β_{limbs} = -0.46 (CI:-0.81,-0.12), t(93)=-2.68, p=0.009
β_{words} = -0.58 (CI:-0.84,-0.32), t(93)=-4.38, p<0.001
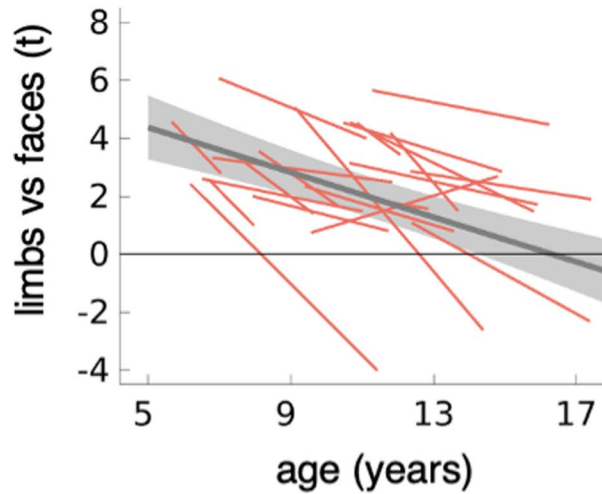
n=21 (96 sessions)

Extended Data Fig. 3 | Developmental changes in word-, face-, and limb-selectivity are also linked in the right hemisphere. (**a**) Limb-selectivity vs face- and word-selectivity in the waning right OTS-limbs. (**b**) Face-selectivity vs. limb- and word-selectivity in the emerging right pFus faces. *Left:* Model prediction for 5–9-year-olds and 13–17-year-olds for the selectivity that defines the ROI as a function of the selectivity to the other two variables. *Middle:* Individual participant data visualized in 3D. In each panel the variable on the z-axis is related to the x- and y-variables. LMM βs, 95%-CIs, t-values, df, and p-values are shown on top. Full statistics are reported in Supplementary Table 11. *Orange arrows:* Individual child data. *Blue arrows:* LMM, same as left panel. *Right:* Rotated version of the plots in the middle column to increase visibility of changes along the horizontal axes. Related to Fig. 4.
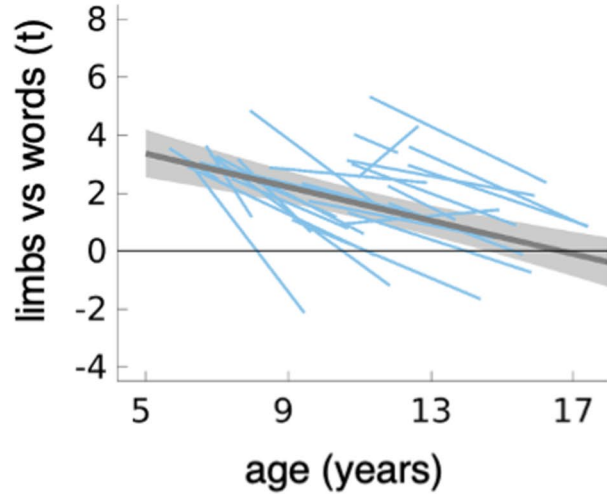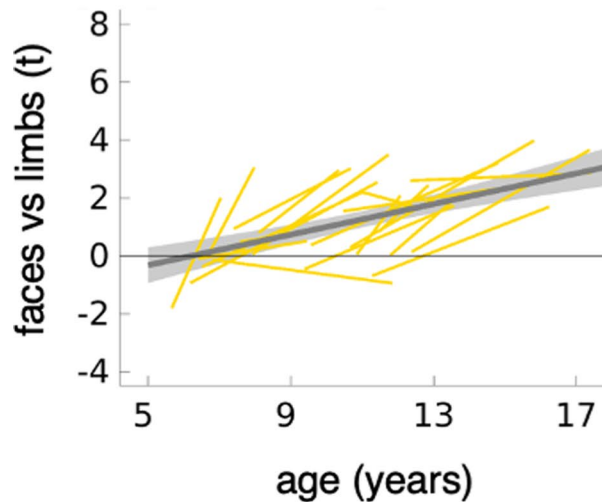
A Waning right OTS-limbs   n=21(100 sessions)

β_{intcpt}=6.3 (CI:4.63,7.98),
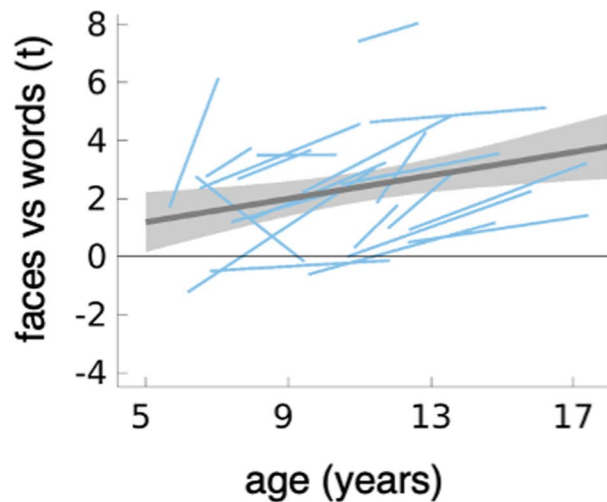t(98)=7.46, p_{FDR}<0.001
β_{age}=-0.03 (CI:-0.04,-0.02),
t(98)=-5.7, p_{FDR}<.001

β_{intcpt}=4.5 (CI:3.23,5.76),
t(98)=7.06, p_{FDR}<0.001
β_{age}=-0.016 (CI:-0.025,-0.007),
t(98)=-3.61, p_{FDR}<0.001

B Emerging right pFus-faces   n=21 (96 sessions)

β_{intcpt}=-1.65 (CI:-2.68,-0.62),
t(94)=-3.19, p_{FDR}=0.003
β_{age}=0.02 (CI:0.015,0.029),
t(94)=5.94, p_{FDR}<0.001

β_{intcpt}=0.18 (CI:-1.52,1.88),
t(94)=0.21, p_{FDR}=0.92
β_{age}=0.017 (CI:0.005,0.029),
t(94)=2.75, p_{FDR}=0.009

Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Pairwise preferences of the ROI-defining category in the right hemisphere. In each plot we show the pairwise preference of the ROI-defining category vs. each of the other two developing categories as a function of age. Thin lines: individual participant data showing the pairwise preference from the initial to end session. Gray line: LMM prediction of pairwise preference based on data from all sessions. Shaded gray: 95%-CI. LMM results (intercept: $\beta_{intercept}$ and slope: β_{age} (rate of change in preference, t/month) and their significance are reported under each panel. (**a**) Waning right OTS-limbs (n = 21 participants, n = 100 sessions). *Left:* Limbs vs faces. *Right:* Limbs vs words. (**b**) Emerging right pFus-faces (n = 21 participants, n = 96 sessions). *Left:* Faces vs limbs. *Right:* Faces vs words. Related to Fig. 5. Statistics in Supplementary Table 15.

nature research

Corresponding author(s):   Kalanit Grill-Spector

Last updated by author(s):   Apr 12, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Code to the 10-category experiment: https://github.com/VPNL/fLoc
Data analysis	Data analysis was performed in MATLAB version 2017b (The MathWorks, Inc.) and using the mrVista software package (https://github.com/vistalab/vistasoft/wiki/mrVista). Quantitative whole brain images of each child and timepoint were processed with the mrQ pipeline (https://github.com/mezera/mrQ). Each participant's individual brain anatomical template was generated using the FreeSurfer Longitudinal pipeline (https://surfer.nmr.mgh.harvard.edu/fswiki/LongitudinalProcessing) using FreeSurfer version 6. Custom code used to generate the main figures will be made available with publication at: https://github.com/VPNL/corticalRecycling.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data required to generate the main figures are available in the GitHub repository (see code availability). Due to the large size of the raw data, it will be made available from the corresponding author upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We collected data from 40 (26 female) children (onset age=5-12 years, M=8.66 years, SD=2.34 years). We selected this age range because (i) it captures the phase in which children start learning to read and (ii) it covers a broad age range spanning childhood and adolescence in which previous studies have documented ventral temporal cortex development. After data exclusions (see below) data from 128 functional sessions of 29 children remained. No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous cross-sectional publications and larger than previous longitudinal studies on VTC development.
Data exclusions	We collected data from 40 (26 female) children. Data from 4 children were excluded because they dropped out of the study after participating only once, and thus did not provide longitudinal data. Data from 7 children were excluded because their data did not pass inclusion criteria (see below). In the remaining 29 children, 29 functional sessions were excluded due to motion, 1 session due to a technical error during acquisition, and 1 session due to aliasing artifacts during acquisition. Therefore, data from 128 functional sessions of 29 children (18 female) are reported in this study. Inclusion criteria were as follows: In each functional session children participated in three runs of the 10 category-experiment. Criteria for inclusion of data were (i) at least 2 runs per session where within-run motion < 2 voxels and between-run motion < 3 voxels, and (ii) at least two fMRI sessions at least six months apart. Because only two of the three runs survived motion quality thresholds for several fMRI sessions, analyses include two runs per child per session to ensure equal amounts of data across participants and sessions.
Replication	We used two independent ways of assessing development of category-selective volume in VTC. In the analysis in Figure 1 an observer-independent approach to assess development of category-selective volume in VTC is used. The results of this analysis are replicated in Figure 2, in which development of category selective volume is measured using the size of manually defined regions of interests.
Randomization	Randomization is not applicable to this study; All statistics are run on the whole sample as this is a longitudinal study that tracks development over time.
Blinding	Blinding was not possible, however all datasets are processed using the same analysis pipeline. This pipeline is automated except for defining regions of interest (VTC ROIs and functional ROIs, see Fig. 2). The results of Fig. 2 were internally replicated in Fig. 1, which uses an observer-independent approach to assess development of category-selective volume in VTC.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	The sample consisted of children with normal or corrected-to-normal vision who were recruited from local schools in the Bay Area. The diversity of the participants reflects the makeup of the Bay Area population. 62.5% of children were Caucasian, 20% were Asian, 5% were Native Hawaiian, 5% were Hispanic, and 7.5% were multiracial or from other racial/ethnic groups. We collected data from 40 (26 female) children (onset age=5-12 years, M=8.66 years, SD=2.34 years). We selected this age

range because (i) it captures the phase in which children start learning to read and (ii) it covers a broad age range spanning childhood and adolescence in which previous studies have documented ventral temporal cortex development.

Recruitment	Participants were recruited from local schools in the Bay Area and/or were children of Stanford staff. The diversity of the participants reflects the makeup of the Bay Area population. No children were home schooled.
Ethics oversight	Institutional Review Board of Stanford University

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type	Block design with 4 seconds block duration
Design specifications	Participants completed two runs of the category experiment. Each run is approx. 5 minutes long.
Behavioral performance measures	Participants were instructed to view the images while fixating on a central dot and to perform an oddball task during scanning. We report both the median performance and standard deviation for the oddball task.

Acquisition

Imaging type(s)	functional, structural
Field strength	3
Sequence & imaging parameters	MRI data were acquired at the Center for Cognitive Neurobiological Imaging at Stanford University on a 3 Tesla GE Discovery MR750 scanner (GE Medical Systems) using a phase-array 32 channel head coil. Whole brain anatomical scans were collected using quantitative MRI (qMRI, Mezer, A. et al. Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. Nat. Med. 19, 1667–1672 (2013)) with a spoiled gradient echo sequence using multiple flip angles (α=4°, 10°, 20°, 30°), TR=14ms and TE=2.4ms. The scan resolution was 0.8x0.8x1.0mm3 (later resampled to 1mm isotropic). For T1-calibration we acquired spin-echo inversion recovery scans with an echo-planar imaging read-out, spectral spatial fat suppression and a slab inversion pulse. These scans were acquired at TR=3s, inplane resolution=2mmx2mm, slice thickness=4mm and 2x acceleration, echo time=minimum full. Functional data were collected using the same scanner and head coil as the structural images. Slices were oriented parallel to the parieto-occipital sulcus. The simultaneous multi-slice, one-shot T2* sensitive gradient echo EPI sequence was acquired with a multiplexing factor of 3 to acquire near whole brain coverage (48 slices), FOV=192mm, TR=1s, TE=30ms, and flip angle=76°. Resolution was 2.4 mm isotropic.
Area of acquisition	Whole brain scans were performed for structural data, near whole brain coverage for functional data.

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	Functional data were analyzed using mrVista (https://github.com/vistalab/vistasoft/wiki/mrVista). Motion correction was performed both within and across functional runs. No spatial smoothing and no slice-timing correction were performed. Time courses were transformed into percentage signal change by dividing each timepoint of each voxel's data by the average response across the entire run.
Normalization	Data were not normalized. Functional data from each session were aligned to the individual subject's across years within-subject template. This within-subject template was created using individual T1 brain volumes from the child's multiple timepoints. Each participant's brain anatomical template was generated using the FreeSurfer Longitudinal pipeline (https://surfer.nmr.mgh.harvard.edu/fswiki/LongitudinalProcessing) using FreeSurfer version 6.0.
Normalization template	No template is used. The data were not normalized to a template. Individual within-subject templates were created and functional data of each participant were aligned to the individual template (see above).
Noise and artifact removal	Motion correction was performed both within and across functional runs.
Volume censoring	N/A

Statistical modeling & inference

Model type and settings	Linear mixed models (LMMs) were used for statistical analyses because (i) the data has a hierarchical structure with sessions being nested within each participant, and (ii) sessions were unevenly distributed across time. Models were fitted using the 'fitlme' function in MATLAB version 2017b (The MathWorks, Inc.). The LMMs were random-intercept models with age as a fixed effect and participant as a random effect.
Effect(s) tested	Using the Linear mixed models we tested if the volume of category-selective activation (Fig. 1,2)/mean selectivity (Fig. 3,5) can be predicted by age (fixed effect). In analyses related to Fig 4. we tested if the selectivity to one category can be

predicted by selectivity to two other categories.

Specify type of analysis: ☐ Whole brain ☒ ROI-based ☐ Both

Anatomical location(s) Anatomical ROIs (ventral temporal cortex) were individually defined using anatomical landmarks. Functional ROIs were individually defined using a threshold of t>3.

Statistic type for inference
(See Eklund et al. 2016) Model parameters and contrasts are computed at the voxel level; there is no cluster size correction

Correction FDR for LMM models

Models & analysis

n/a	Involved in the study
☒ ☐	Functional and/or effective connectivity
☒ ☐	Graph analysis
☒ ☐	Multivariate modeling or predictive analysis