# Title: Cortical recycling in high-level visual cortex during childhood development

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

**Affiliations:**

[1] Department of Psychology, Stanford University, Stanford, CA.

[2] Neurosciences Program, Stanford University, Stanford, CA.

[3] Princeton Neuroscience Institute, Princeton University, NJ.

[4] Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA.

*Correspondence to: kalanit@stanford.edu

## Summary

Human ventral temporal cortex (VTC) contains category-selective regions that respond preferentially to ecologically-relevant categories such as faces[1], bodies[2], places[3], and words[4] and are causally involved in the perception of these categories[5–7]. However, it is unknown how these regions develop during childhood. Here we used functional MRI and images from many categories to measure longitudinal development of category-selectivity in individual school-age children over the course of 5 years. We show that from young childhood to the teens, face- and word-selective regions in VTC expand and increase in their respective category-selectivity, but limb-selective regions in VTC shrink and lose their preference for limbs. Critically, as a child develops, increases in their face- and word-selectivity are directly linked to decreases in limb-selectivity. These data show that during childhood limb-selectivity in VTC is repurposed into word- and face-selectivity providing the first empirical evidence for cortical recycling[8] during childhood development. These results suggest a rethinking of prevailing hypotheses that cortical development involves sculpting of new representations upon general-purpose cortex[9,10]. Instead, they suggest a new hypothesis that during development VTC representations adjust to changes in the salience and social relevance of visual inputs[11], which has important implications for both typical and atypical brain development.

A central question in neuroscience is how does cortical function develop? 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 behavior 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[12–14]. However, experience with other categories, such as words, doesn't begin until later in childhood when children learn to read. Because face and word recognition require fine visual acuity afforded by foveal representations in lateral VTC[15], two main developmental theories have been proposed. The theory of functional refinement predicts that category-selective regions emerge from raw, general-purpose cortex[9,10] that has some basic property[16], such as foveal bias[17,18]. The theory of competition posits that rivalry for cortical resources may constrain development[8,19], leading to recycling[8] of cortex selective to one category earlier in childhood (e.g., faces) to be selective to other stimuli (e.g., words) with later demands for reading[8,20]. This intense theoretical debate can only be addressed by longitudinal brain measurements in school-age children, because reading acquisition is specific to humans, and the fusiform gyrus, which is the anatomical structure where face-selective regions reside, is hominoid-specific.

Here, we addressed this key gap in knowledge using longitudinal fMRI in 29 children (initially 5–12 years old) to measure development of many category representations and their relationship. Children were scanned repeatedly over the course of 1 to 5 years (mean±SD: 3.75±1.5 years, **Fig. S1A**), with an average of 4.4±1.92 fMRI sessions per child and 128 included sessions overall (Methods). During fMRI, children viewed images from 10 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, **Fig. S1D**). 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 standard brain space[21].

**How does selectivity in VTC develop?**

To assess development of VTC, we first quantified the volume of category-selective activation in VTC as a function of age. Category-selectivity was computed by contrasting responses to each category vs. all other categories except the other category from the same domain (e.g., limbs vs. all other categories except bodies, $t>3$, voxel level). VTC was anatomically defined on the cortical surfaces of each child and divided into lateral and medial sections (**Fig. 1B**). This division captures the center-periphery organization of VTC[15,18], where lateral VTC represents the central visual field, and medial VTC represents the periphery. To test for age-related changes in the volume of category-selective activation, we used linear mixed models (LMMs) with age as a fixed effect and participant as a random effect. LMMs with intercepts (indicating initial volume) that varied across participants and a constant slope (indicating the rate of change in volume with age, **Fig 1A,C**) fit the data best in the majority of cases (Methods).

Results reveal differential development of category-selectivity across VTC partitions and categories: (1) Significant development of category-selectivity occurred in lateral, but not medial VTC (**Fig. 1C, Fig. S2, Table S1-2**). (2) Surprisingly, there were both developmental increases and decreases in the volume of category-selective activation in lateral VTC: word- and face-selective activation increased, but limb-selective activation decreased (**Fig. 1A,C, Fig. S2,** full statistics in **Table S1**).

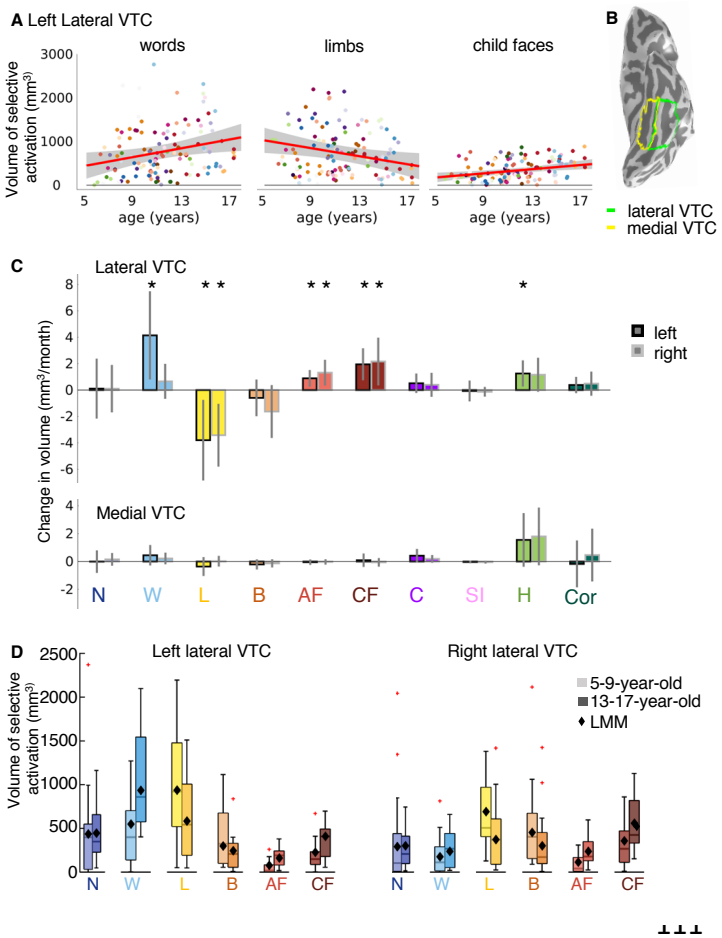


**Fig. 1. Developmental increases and decreases in category-selective activation in lateral VTC.** **(A)** Volume of word-, limb- and child-face-selective activation by age. Each dot is a session and colored by participants. *Red line:* LMM prediction of category-selective activation by age. Shaded *gray:* 95% confidence interval (CI). **(B)** Lateral and medial VTC on inflated cortical surface of a 5-year-old. **(C)** LMM slopes: change in category-selective activation per month. *Error bars:* 95% CI. *Asterisks*: significant after FDR-correction ($p<0.05$). All scatterplots in **Fig. S2**. **(D)** *Boxplots:* average category-selective activation by age group. One session per child is included per boxplot. *Diamonds:* average selective volume predicted by LMM. *Crosses:* outliers; All categories in **Fig. S5**. *Acronyms:* N: numbers; W: words; L: limbs; B: bodies; AF: adult faces; CF: child faces, C: Cars, SI: String instruments, H: Houses, Cor: Corridors.

Interestingly, during childhood, the volume of word-selective activation significantly increased in left, but not right, lateral VTC (**Fig. 1C,D Fig. S2, Table S1**), while number-selective activation remained unchanged (**Fig. 1C, Fig. S2, Table S1**). Examining the average volume of activations (**Fig. 1D-boxplots**) as well as LMM predictions (**Fig 1D-diamonds**), revealed that word-selectivity doubled on average from ~500 $mm^3$ in 5–9-year-olds to ~1000 $mm^3$ in 13–17-year-olds. Notably, as word-selective activation doubled, volume of limb-selective activation halved (**Fig. 1D**). In fact, limb-selective activation significantly decreased bilaterally, but body-selective activation remained stable (**Fig. 1C**). Additionally, the volume of face-selective activation increased bilaterally for both adult faces and child faces **(Fig. 1C,D)**. Similar

longitudinal development in lateral VTC was observed for other contrasts (**Fig. S3**) and for other metrics, such as selectivity within constant-sized regions (**Fig. S4**).

## Is development anatomically specific?

To determine the anatomical specificity of the observed development, we defined limb-, word-, and face-selective regions of interest (ROIs) in each participant and session (**Fig. 2A-C,** t-value > 3 voxel-level) and examined them longitudinally. The location of category-selective ROIs remained largely the same as ROIs expanded or shrank across childhood (**Fig. 2A-C**). The growth of word- and face-selective regions was anatomically specific: posterior but not anterior ROIs significantly expanded (**Fig. 2D, Table S3**). Activation for words in the left posterior occipitotemporal sulcus (pOTS-words) grew significantly, but not in the mid occipitotemporal sulcus (mOTS-words). Bilateral activation for faces in the posterior fusiform (pFus-faces) grew significantly, but not in the mid fusiform (mFus-faces). In contrast to the growth of pOTS-words and pFus-faces, OTS-limbs shrank significantly in both hemispheres (**Fig. 2D, Table S3**).

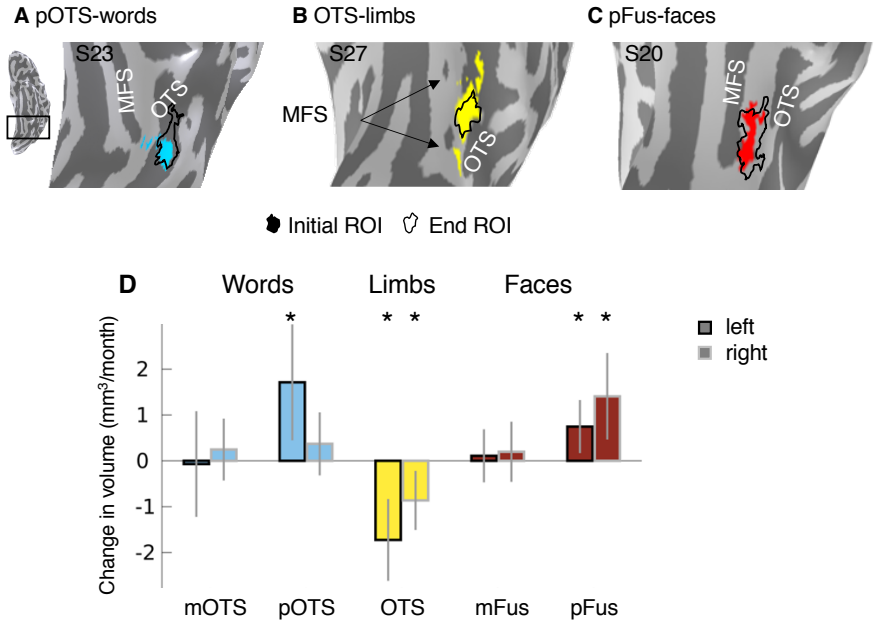


**Fig. 2. Development of category-selective ROIs.** Initial ROIs (colored) and end ROIs (outline) in 3 example children: **(A)** left pOTS-words at age 10 and 15, **(B)** left OTS-limbs at age 11 and 13, **(C)** left pFus-faces at age 9 and 14. *MFS:* mid fusiform sulcus; *OTS:* occipito-temporal sulcus. **(D)** LMM slopes: change in volume of category-selective regions per month. *Error bars:* 95% CI; *Asterisks:* significant development, p<0.05, FDR-corrected.

**What changes in emerging and waning ROIs?**

We next assessed the functional properties of the developing regions. 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 left lateralization of the development of word-selectivity (right hemisphere-**Fig. S7)**. Since developing regions are not completely independent from the original ROIs, we repeated the analysis in independent ring-shaped ROIs centered on initial functional ROIs, yielding similar findings (**Fig. S8**).

In the emerging left pOTS-words, selectivity to words significantly increased with age, as expected (LMM of selectivity, age: continuous fixed effect; participant: random effect). At the same time, selectivity to limbs and bodies significantly decreased (**Fig. 3A-left, Table S4**), but there were no changes in selectivity to other categories (**Fig. S7**, **Table S4**). To elucidate if 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 to each category as a function of age (age: continuous fixed effect; participant: random effect). Results show that developmental changes in word-selectivity were associated with significant increases in the responses to words with no significant changes to other categories except that responses to string instruments also significantly increased (**Table S5)**. Indeed, pOTS-responses to words are higher in teens (13-17-year-olds) than children (5-9-year-olds, **Fig. 3A-right, Table S5**) in correspondence with LMM predictions.

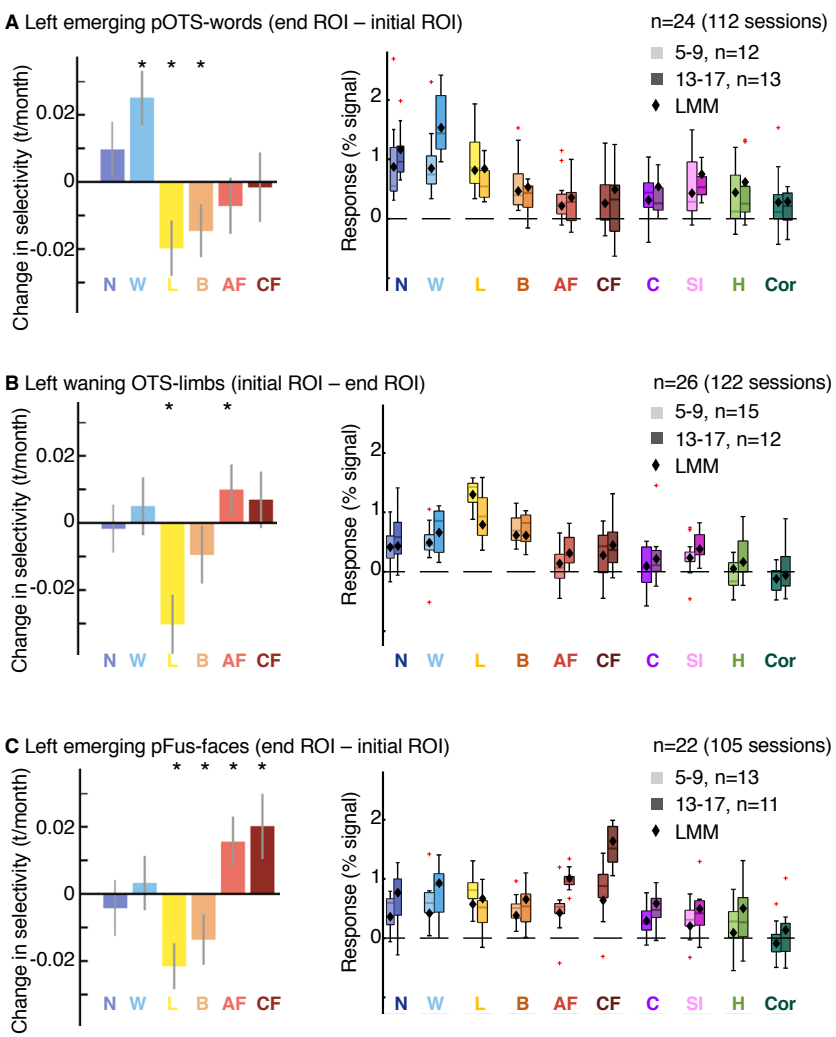


**Fig. 3. Age-related increases in word- and face-selectivity parallel decreases in limb-selectivity in the developing regions**. *Left:* changes in selectivity by age (LMM slopes). *Error bars: 95% CI; Asterisks:* significant development: $p<0.04$, FDR corrected*;* Data for all categories in **Fig. S7**. *Right*: *Boxplots*: average response amplitudes for 5-9-year-olds and 13-17-year-olds for the 10 categories. One functional session per child is included per boxplot. *Diamonds:* estimated response amplitudes from LMM; *Crosses:* outliers. **(A)** Emerging word ROI. **(B)** Waning limb ROI. **(C)** Emerging face ROI.

Similarly, in the emerging pFus-faces, selectivity to faces increased (**Fig. 3C**, **Fig. S7**). At the same time, selectivity to limbs decreased bilaterally and selectivity to bodies decreased in the left hemisphere (**Fig. 3C**). There were no significant changes in selectivity to other categories (**Fig. S7, Table S4**). Increases in selectivity to faces were associated with significant increases in responses to faces (**Table S5**). Indeed, 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, and

there was a trend for an increase in responses to cars (**Fig. 3C-right, Table S5**). Thus, developmental increases in word and face selectivity in emerging word and face ROIs, respectively, are driven by increased responses to their preferred category, rather than decreased responses to nonpreferred categories.

As OTS-limbs is located between pOTS-words and pFus-faces, we asked if increased responses to faces and words also occur in the waning OTS-limbs. We found that not only did responses to adult faces significantly increase with age, but also responses to limbs significantly decrease with age (**Fig. 3B-right, Table S5, Fig. S7**). 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. Responses to other categories remained stable (**Fig. 3B-right, Fig. S7**). These changes in responses resulted in significant decreases in limb-selectivity (**Fig. 3B-left**, **Fig. S7),** significant increases in face-selectivity (**Fig. 3B-left**), and no changes in selectivity to other categories including bodies (**Fig. S7**). Therefore, developmental decreases in limb-selectivity of the waning OTS-limbs reflect decreased responses to the preferred category together with increased responses to faces.

As we found developmental decreases in limb-selectivity in both emerging and waning ROIs, we tested if this is a general phenomenon across lateral VTC. Analyses of lateral VTC excluding voxels that were selective in the first sessions to categories showing development, revealed no significant decreases in limb-selectivity in the remainder of lateral VTC (**Fig. S9**). This suggests that developmental changes in limb-selectivity are most prominent in the emerging and waning ROIs.

**Are changes in selectivity linked?**

We tested if there is a quantitative relationship between selectivities to faces, words, and limbs in the developing ROIs. Model comparison 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 $\chi^2 \geq 4.76$, $p \leq 0.029$, **Table S6**). Moreover, in all developing ROIs, selectivity to the preferred category was significantly and negatively related to selectivity to the other two categories (LMM predictors (bs): **Fig. 4, Fig. S10, Table S7**). E.g., in the emerging pOTS-words, higher word-selectivity is significantly linked with both lower face- and limb-selectivity (**Fig. 4A, Table S7**).

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**). In the emerging pOTS-words, 5-9-year-olds have positive selectivity to limbs and negative selectivity to faces. By age 13-17 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 (**Table S8**). Notably, this developmental pattern is visible in individual children from their initial age (**Fig. 4A-right**, arrow bases) to their end age (**Fig. 4A-right**, arrow heads). Similarly, in the emerging pFus-faces, 5-9-year-olds have positive limb-selectivity and mild word-selectivity (**Fig. 4C**). By age 13-17, as face-selectivity has increased, limb-selectivity is lost and there is little change to word-selectivity (**Fig. 4C Fig. S10, Table S8**).

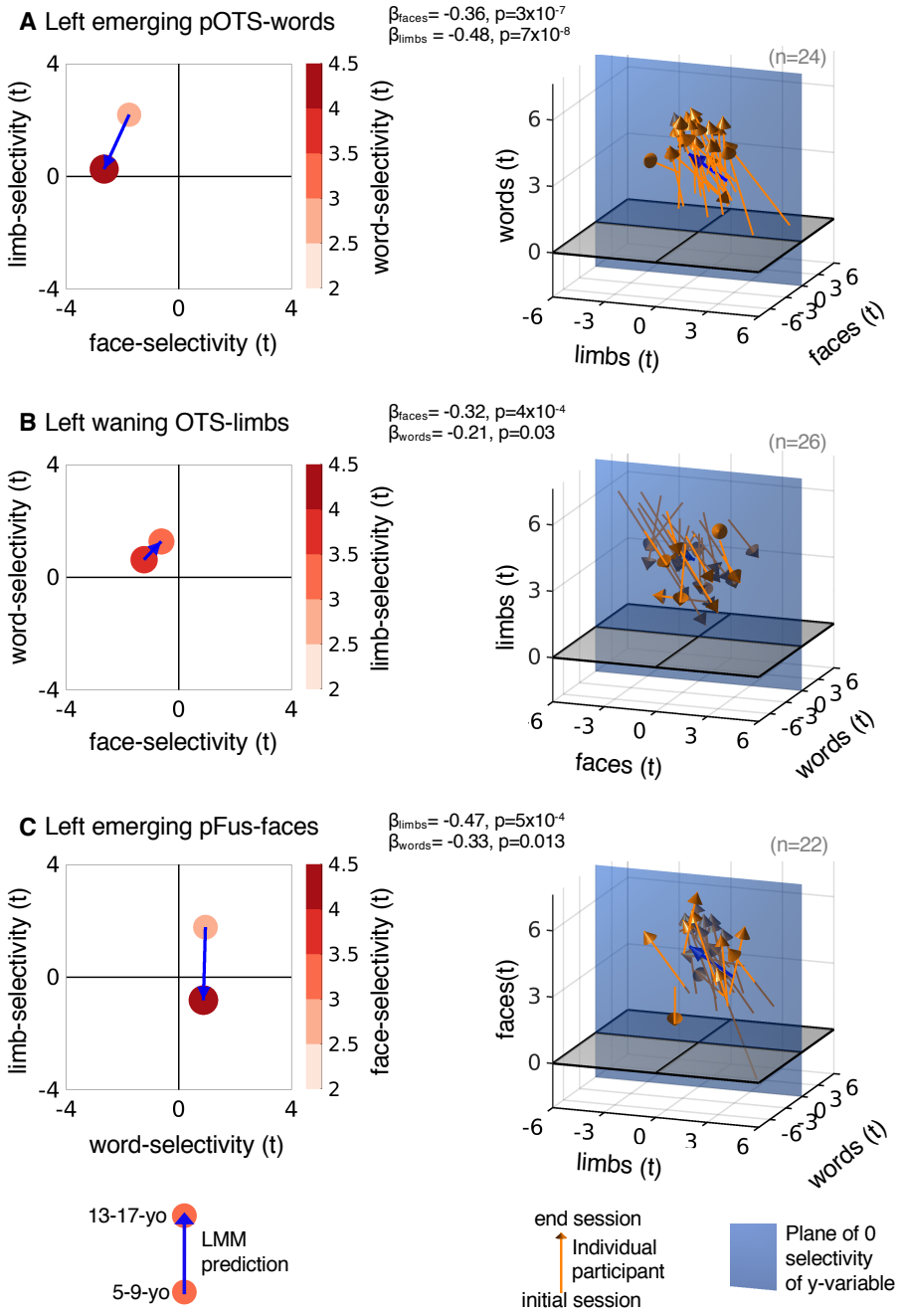


**Fig. 4. Developmental changes in word-, face-, and limb-selectivity are linked.**
*Left:* LMM prediction of category selectivity (circle, see colorbar) to words **(A)**, limbs **(B)**, and faces **(C)** vs. selectivity to the other two categories in 5-9-year-olds and 13-17-year-olds. *Right:* Individual child data. *Orange arrows:* negative selectivity in y-variable, *brown arrows:* positive selectivity in y-variable. *Blue arrows*: LMM, same as in corresponding left panel.

285

286 In the waning OTS-limbs, 5-9-year-olds exhibit negative face-selectivity and mild word-
287 selectivity (**Fig. 4B-left**). As limb-selectivity declines by age 13-17, both word- and face-
288 selectivity increase (**Fig. 4B**-left). While limb-selectivity consistently declined across individuals,

289 there was more variability in individual developmental trajectories compared to the emerging ROIs
290 (**Fig. 4B-right**). In some children limb-selectivity was replaced by word-selectivity and in others
291 with face-selectivity.

292

293 **Discussion**

294 Our longitudinal measurements in children using a large range of ecologically-relevant
295 categories reveal three new insights about the functional development of high-level visual cortex.
296 First, we find that childhood development is not only associated with growth of category-selective
297 regions and increases in selectivity[9,10,22], but also involves loss of selectivity. Second, teens' brains
298 are not just more sculpted versions of children's brains. In contrast to the prevailing view
299 suggesting that children's VTC is indistinctive[9,10,19], young children's VTC is actually more
300 selective for limbs than it is later in childhood. Early selectivity to limbs may be related to the
301 visual salience of limbs early in development, as toddlers look at hands more than faces[11] when
302 they learn to gesture[23] and manipulate objects. Third, our results provide the first empirical
303 evidence for recycling[8] of category-selectivity in high-level visual cortex during childhood.
304 However, contrary to previous predictions that face-selectivity is recycled to word-selectivity[8,19]
305 during development, our results show that limb-selectivity is recycled to both word- and face-
306 selectivity. This recycling occurs via a mechanism of decreasing responses to limbs and increasing
307 responses to both faces and words. This recycling is specific to limb-selectivity as body-selectivity
308 remains unchanged, consistent with previous findings[10,24].

309 Our results require a rethinking of prevailing developmental theories that propose that
310 cortical development involves sculpting of new representations upon general-purpose cortex[9,10].
311 Instead, our data suggest that during childhood, cortical selectivity can change from one category

to another. 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, our data suggest 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, such as faces and words.

These findings have important implications for understanding typical[25] and atypical[26–29] 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 related to reading[30] and social perception[26,27]. Second, cortical recycling may be particularly profound in cases of childhood visual deprivation[28] or brain lesions[29]. It will be important to determine if there is a critical period during development in which cortical recycling occurs, or if it occurs throughout the lifespan[31]. Finally, as deep convolutional neural networks predict VTC responses[32,33], future computational modeling could reveal insights to which computational constraints yield both developmental increases and decreases in category selectivity.

## 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. Parvizi, J. *et al.* Electrical stimulation of human fusiform face-selective regions distorts face perception. *J. Neurosci.* **32**, 14915–14920 (2012).
6. Gaillard, R. *et al.* Direct Intracranial, fMRI, and Lesion Evidence for the Causal Role of Left Inferotemporal Cortex in Reading. *Neuron* **50**, 191–204 (2006).
7. Mégevand, P. *et al.* Seeing scenes: Topographic visual hallucinations evoked by direct electrical stimulation of the parahippocampal place area. *J. Neurosci.* **34**, 5399–5405 (2014).
8. 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).
9. Golarai, G. *et al.* Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nat. Neurosci.* **10**, 512–522 (2007).
10. 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).
11. 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).
12. Deen, B. *et al.* Organization of high-level visual cortex in human infants. *Nat. Commun.* **8**, 13995 (2017).
13. de Heering, A. & Rossion, B. Rapid categorization of natural face images in the infant right hemisphere. *Elife* **4**, 1–14 (2015).
14. Livingstone, M. S. *et al.* Development of the macaque face-patch system. *Nat. Commun.* **8**, 1–12 (2017).
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. Srihasam, K., Vincent, J. L. & Livingstone, M. S. Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nat. Neurosci.* **17**, 1776–1783 (2014).
17. 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**, 1–12 (2018).

18. Nordt, M. *et al.* Learning to Read Increases the Informativeness of Distributed Ventral Temporal Responses. *Cereb. Cortex* 1–16 (2018). doi:10.1093/cercor/bhy178

19. Behrmann, M. & Plaut, D. C. A vision of graded hemispheric specialization. *Ann. N. Y. Acad. Sci.* **1359**, 30–46 (2015).

20. 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).

21. Frost, M. A. & Goebel, R. Measuring structural-functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage* **59**, 1369–1381 (2012).

22. 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**, (2007).

23. Liszkowski, U., Carpenter, M. & Tomasello, M. Pointing out new news, old news, and absent referents at 12 months of age. *Dev. Sci.* **10**, 1–7 (2007).

24. 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).

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

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

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

28. 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).

29. 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).

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

31. 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).

32. 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.* **2018-Decem**, 8388–8399 (2018).

33. Khaligh-Razavi, S. M. & Kriegeskorte, N. Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. *PLoS Comput. Biol.* **10**, (2014).

## Methods

### Participants

Children with normal or corrected-to-normal vision 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. Prior to the start of the study, parents gave written consent, and children gave written assent. The study was approved by the Institutional Review Board of Stanford University.

Prior to their first MRI session, children were trained in a scanner simulator to acclimate them to the scanner environment and to enhance quality of MRI data. In the simulator, children practiced laying 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 (26 female) children (onset age=5-12 years, M=8.66 years, SD=2.34 years, **Fig. S1)**. 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[10,22].

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 neurotypical children (18 female, 11 male) are reported in this study (**Fig. S1A,B** has an

overview of the included and excluded sessions). Initial ages of the included children ranged from 5 to 12 years (mean=9.19, SD=2.13).

Participants were scanned using functional and structural MRI for 1 to 5 years. When possible, children participated in 1 to 2 functional scans per year. Additionally, children participated in 1 structural MRI session per year. Each child participated in at least 2 and up to 10 fMRI sessions (mean=4.41, SD=1.92) with the time interval between the first and last fMRI scan ranging from 10 months to 5 years (mean=45 months, SD=18 months, **Fig. S1C**). 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 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[34]) with a spoiled gradient echo sequence using multiple flip angles ($\alpha$=4°, 10°, 20°, 30°), TR=14ms and TE=2.4ms. The scan resolution was 0.8x0.8x1.0mm$^3$ (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=2x2mm$^2$, slice thickness=4mm and 2x acceleration, echo time=minimum full.

452 *Functional imaging*

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

458

459 *10 category experiment*

460 Participants completed three runs of a 10-category experiment[17,18,35]. During each run
461 participants viewed images from five domains each comprising images of two categories including
462 faces (adult faces, child faces), body parts (headless bodies, limbs), objects (cars, string
463 instruments), places (corridors, houses) and characters (pseudowords, numbers). Following prior
464 work on visual representations, we define a visual category as a set of exemplars sharing the same
465 parts and configuration, e.g., limbs[36–39] and domain as a grouping of one or more categories that
466 share semantic association (whether or not they share visual features) and are thought to require
467 distinct processing mechanisms[40] e.g., houses and corridors are both places, but are visually
468 dissimilar. Examples of stimuli are shown in **Fig. S1D**.

469 Images were presented in 4 s blocks, at a rate of 2 Hz and did not repeat across the course of
470 the experiment. Image blocks were intermixed with gray luminance screen baseline blocks. Blocks
471 were counterbalanced across categories and baseline. Stimuli were grayscale and contained a
472 phase-scrambled background generated from randomly selected images. Participants were
473 instructed to view the images while fixating on a central dot and to perform an oddball task.

Participants pressed a button whenever an image comprising only the phase-scrambled background appeared.

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).

*Inclusion criteria*

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. For sessions with all 3 runs passing motion quality criteria, 2 runs with lowest within-run motion were included.

*Structural MRI data analysis and individual template creation*

Quantitative whole brain images of each child and timepoint were processed with the mrQ pipeline (https://github.com/mezera/mrQ[34]) to generate synthetic T1 brain volumes. For each child, the synthetic T1 brain volumes from their multiple timepoints were used to generate a within-subject brain volume template. Each participant's brain anatomical template was generated using the FreeSurfer Longitudinal pipeline (https://surfer.nmr.mgh.harvard.edu/fswiki/LongitudinalProcessing[41]) using FreeSurfer version

6.0. The gray-white matter segmentation of each participant's within-subject brain template was manually edited to fix segmentation errors (e.g., holes and handles) to generate an accurate cortical surface reconstruction of each participant's brain.

The motivation for aligning the functional data to the within-subject-template were (i) to enable comparison of regions of interest (ROIs) from different timepoints in the same brain volume for each participant and to (ii) minimize potential biases which can occur from aligning longitudinal data to the anatomical volume from a single timepoint[41]. On average 2.48 (SD=0.69) synthetic T1s were used to generate the within-subject-template (min=2, max=5). Functional data from all sessions of a participant were aligned to their within-subject brain template. In 17 participants the last fMRI session that was included was conducted after the within-subject 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-subject-template (excluding 2 participants whose last T1 could not be used because of technical error during acquisition and subject motion).

*Definition of anatomical regions of interest (lateral and medial VTC ROIs)*

On the inflated surface of each hemisphere in each participant, we defined anatomical regions of interest (ROIs) of the lateral and medial VTC (**Fig. 1B**) as in previous publications [18]. We first defined the VTC anatomically and then separated it to 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-subject template. 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. To estimate the contribution of each of the 10 conditions (corresponding to the 10 image categories) a general linear model (GLM) was fit to each voxel by convolving the stimulus presentation design with the hemodynamic 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-value > 3 (voxel-level) for the contrast of interest. We also performed a complementary threshold-independent analysis in constant-sized regions (see *Control analysis* below). Category contrasts (**Fig. 1**) were computed by contrasting responses to each category with all other categories from all other domains (i.e., words vs. all other categories except numbers). We also computed contrasts for domains (faces, body parts, characters, objects, places), see supplemental analyses (**Fig. S3**). For domain contrasts, responses to each domain were contrasted with all other domains (i.e., characters vs. all others).

543

544 *Definition of functional regions of interest*

545 To examine the anatomical specificity of the observed development of selectivity (**Fig. 1**),
546 we defined word-, limb-, and face-selective functional regions in each participant (see **Fig. 2A-C**).
547 For the definition of functional ROIs, a threshold of a t-value > 3 (voxel-level) was used. All ROIs
548 were defined in each participant's native cortical surface generated from the within-subject brain
549 template. *Word-selective regions* were defined as above-threshold clusters for the contrast words
550 vs. all categories except numbers that straddled the occipito-temporal-sulcus (OTS). The more
551 anterior cluster was defined as mOTS-words and the posterior cluster as pOTS-words. These
552 clusters are also referred to as the visual word form area (VWFA 1 and 2, respectively[4]). The *limb-*
553 *selective region* was defined as above-threshold cluster of voxels for the contrast limbs vs. all
554 categories except bodies straddling the OTS and was labelled OTS-limbs. This cluster is also
555 referred to as the fusiform body area (FBA[2]). *Face-selective regions* were defined using the
556 contrast faces (adult and child) vs. all other categories, because we observed similar development
557 for both face types in the analysis related to **Fig. 1**. Face-selective clusters were defined as above-
558 threshold clusters on the lateral fusiform gyrus. The more anterior cluster typically aligns with the
559 anterior tip of the MFS, and was defined as mFus-faces, while the more posterior cluster was
560 defined as pFus-faces[42]. These two face-selective clusters are also referred to as the fusiform face
561 area (FFA 1 and 2, respectively[1]).

562 For supplemental analyses (**Fig. S6**) we also defined place-, character-, and combined body-
563 part-selective regions. Place-selective regions were defined as above threshold clusters on the
564 collateral sulcus that respond to houses and corridors vs. all other categories. This region is also
565 referred to as parahippocampal place area (PPA[3]). Character-selective regions were defined as

566 above-threshold clusters on the OTS that responded more strongly to words and numbers vs. all
567 other categories. Similarly, a body-part-selective region was defined as above-threshold clusters
568 on the OTS that responded more strongly to bodies and limbs than the other categories.
569
570 *Emerging and waning ROIs*
571 To characterize the selectivity and responses within VTC regions that changed with
572 development, we defined in each participant's VTC ROIs that: (i) either gained selectivity to faces
573 or words or (ii) lost selectivity to limbs during childhood. To do so, we defined emerging face and
574 word ROIs as well as waning limb ROIs. For words and faces, where selective voxels emerge
575 during development, the emerging ROI is the region that was selective to faces (or words) in the
576 end timepoint but was not selective to that category in the initial timepoint (**Fig. 2A,C**). We call
577 the region that is the difference between the end and initial ROIs the *emerging* pOTS-words and
578 *emerging* pFus-faces. For limbs, were selective voxels decline during development, we call the
579 difference between the initial and end ROIs the *waning* OTS-limbs. That is, the waning OTS-limbs
580 is the region that was within the ROI in the initial timepoint but not in the end timepoint (**Fig. 2B**).
581 For a given participant, the initial ROI corresponds to the first fMRI session in which the ROI
582 could be identified, and end ROI corresponds to the last fMRI session in which the ROI could be
583 identified.
584
585 *Control analysis of selectivity development in independent ring-shaped ROIs*
586 The emerging and waning ROIs were defined from individual participant's functional ROIs
587 on their native brain anatomy and thus capture precisely the part of cortex that undergoes
588 development. As these ROIs are not completely independent from the original ROIs, we sought to

validate these results in an independent manner. Thus, we conducted a complementary analysis to evaluate responses in independent ring ROIs. First, we created two disk ROIs centered at the center coordinate of the initial functional ROI. One ROI was sized to match the surface area of the initial ROI, and the other, sized to match the corresponding end ROI. Then, we defined the area in between these two disk ROIs as the independent the ring-shaped ROIs.

Within each ring-shaped ROI, we measured responses to the 10 categories as well as selectivity to each category (**Fig. S8**). Importantly, this analysis replicated the significant decrease in limb-selectivity in the 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 figure legend of **Fig. S8** for statistics). In left pOTS-words, we find a similar trend for increasing word-selectivity (**Fig. S8A**) and in right pFus-faces we find a similar trend for increasing face-selectivity (**Fig. S8D,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: estimating changes in selectivity across development 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 ROI that significantly responded to one category (vs. all other categories except the other category from the same domain, threshold of $t>3$, voxel-level). While this threshold guarantees significant selectivity, ultimately the threshold value is a number decided by the experimenter.

612 To ensure that findings of developmental effects do not depend on the threshold used, we
613 performed a complementary analysis of category-selectivity in lateral VTC across development
614 which did not depend on a threshold. Across development, we kept the number of voxels within
615 VTC constant and evaluated their mean selectivity (t-value) to a category. Specifically, at each
616 timepoint we selected the 20% most selective voxels (i.e., the voxels with the highest t-values) for
617 each category contrast and calculated their mean t-value (**Fig. S4**). The 20% most selective voxels
618 are determined in each session independently, to avoid biasing the selection of voxels to a specific
619 timepoint. Then we used LMMs to determine if the mean selectivity of these voxels changes over
620 time. LMMs can be expressed as: *t-value ~ age in months + (1|participant).*
621
622 Results of this analysis largely replicate the main finding presented in **Fig 1C,** and reveal (i)
623 a significant increase in word-selectivity in the left hemisphere (see legend of **Fig S6** for statistics),
624 (ii) a bilateral decrease in limb-selectivity, (iii) an increase in selectivity to adult faces in the right
625 hemisphere, (iv) a bilateral increase in selectivity to child faces and (v) an increase in selectivity
626 to houses in the left hemisphere.
627
628 *Is the decrease in limb-selectivity uniform across lateral VTC?*
629 As we found accumulating evidence for reductions in limb-selectivity in lateral VTC, an open
630 question is whether the decrease in limb-selectivity occurs across the entire lateral VTC or whether
631 it is restricted to the regions in which selectivity changes across development. To this end, we
632 aimed to assess development of limb-selectivity in VTC excluding voxels to any category showing
633 significant development (**Fig. 1**). Thus, we identified in each participant's first session the voxels
634 selective for these stimuli, and then measured limb-selectivity across development in the remaining

lateral VTC voxels. Next, we used LMMs to test if there was a significant decrease in the mean selectivity to limbs in these remaining lateral VTC voxels (**Fig. S9**).

While there was a trend for a decrease in selectivity in both hemispheres, the effects were not significant after correcting for multiple comparisons (left: slope=-0.0039 t/month, p=0.16; right: slope=-0.0036 t/month, p=0.16, FDR corrected). These results suggest that the decrease in limb-selectivity appears to be strongest in the developing category-selective regions.

*Statistical Analyses*

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 (**Fig. S1A**). Models were fitted using the '*fitlme'* function in MATLAB version 2017b (The 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.

LMMs related to **Figure 1** can be expressed as: *volume of selective activation for a category in* $mm^3$ *~ 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 data set including 128 sessions.

Boxplots in **Fig. 1D** showing subsets of the data are used for estimating volume of selective activation in different age groups and visualization purposes, but not to evaluate statistical

significance. One session per child per age group is included in the boxplot. To confirm the validity of the grouping of the age groups for the boxplots in **Fig. 1D** (5-9yo and 13-17yo), we used the LMMs (shown in **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-selective ROIs (**Fig. 2D, Fig. S6**) we used LMMs specified as: *ROI size in* $mm^3$ *~ 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 modeled as *t-value ~ age in months + (1|participant),* and (ii) changes in responses across development (**Fig. 3**-right) as % signal (b)~ *age in months + (1|participant).* Percent signal refers to the change in response for a certain condition relative to 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 fit on the average value in an ROI. Boxplots in **Fig. 3** show mean responses for each of the 10 categories in 5-9-year-olds and 13-17-year-olds. They 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 prediction for the mean age in years of participants in each age group (compare black diamonds to measured median).

LMM analyses related to **Fig. 4** tested if changes in limb-, face-, and word-selectivity in the developing ROIs were related to each other. For each emerging and waning ROI, we tested if selectivity for one category (e.g., word-selectivity in the emerging pOTS-words) was predicted by

selectivity to the other categories (e.g., limb- and face-selectivity in the emerging pOTS-words). Critically, we tested if an LMM with two predictors (e.g., predicting word-selectivity from both limb- and face-selectivity) is a better model than an LMM with one predictor (e.g., 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 supplemental **Tables S6-7.** Tables are grouped by analysis type.

False-discovery rate (FDR) correction following the procedure by Benjamini and Hochberg[43] as implemented in MATLAB version 2017b (The MathWorks, Inc.) was applied to all p-values to correct for multiple comparisons related to the same analysis.

*Data Availability*

Data and code will be made available with publication at: https://github.com/VPNL/Recycling.

**Methods-only references**

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

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

36. 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).

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

38. 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).

39. 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).

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

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

42. 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).

43. 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).

## Acknowledgements

We thank Laura Villalobos, Erica Yeawon, Hwang, Savana Huskins, Alema Fitisemanu, and Philip Eykamp for manually editing gray-white matter brain segmentations. We thank Caitlyn Estrada and Nancy Lopez-Alvarez for help with data collection, and Rachel Hinds for help with data entry and management. We thank Jon Winawer for fruitful discussions. **Funding:** This work was supported by a fellowship of the German National Academic Foundation awarded to MN (NO 1448/1-1); NIH grant 2RO1 EY 022318 to KGS. NIH training grant 5T32EY020485 (VN); NSF Graduate Research Development Program (DGE-114747) and Ruth L. Kirschstein National Research Service Award (F31EY027201) to JG.

**Author contributions:** M.N. collected data, developed and coded the analysis pipeline, analyzed the data, and wrote the manuscript. V.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. KGS designed the experiment, contributed to the analysis pipeline and data analysis, and wrote the manuscript.

**Competing interests:** Authors declare no competing interests.

**Materials & Correspondence**: Correspondence and material requests should be addressed to kalanit@stanford.edu