Report

Current Biology

Decoding sound content in the early visual cortex of aphantasic participants

Highlights

- Sound decoding is reduced in the early visual cortex of participants with aphantasia
- Across 3 datasets, feedback to the early visual cortex is more general than imagery
- Reduced connectivity in aphantasia may be specific to the early visual cortex
- The trend of higher decoding in the periphery decreases with hierarchy from V1 to V3

Authors

Belén M. Montabes de la Cruz, Clement Abbatecola, Roberto S. Luciani, ..., Petra Vetter, Lucy S. Petro, Lars F. Muckli

Correspondence

lars.muckli@glasgow.ac.uk

In brief

Natural sounds can be decoded from the early visual cortex of blindfolded and congenitally blind participants. Montabes de la Cruz et al. replicate the paradigm with aphantasic participants and report decreased sound decoding, consistent with less-informative top-down projections to the early visual cortex in this population.

Montabes de la Cruz et al., 2024, Current Biology *34*, 1–7
November 4, 2024
https://doi.org/10.1016/j.cub.2024.09.008

CellPress

Please cite this article in press as: Montabes de la Cruz et al., Decoding sound content in the early visual cortex of aphantasic participants, Current Biology (2024), https://doi.org/10.1016/j.cub.2024.09.008

Current Biology

CellPress
OPEN ACCESS

Report

Decoding sound content in the early visual cortex of aphantasic participants

Belén M. Montabes de la Cruz,[1,6] Clement Abbatecola,[1,2,6] Roberto S. Luciani,[1,3] Angus T. Paton,[1,2] Johanna Bergmann,[4] Petra Vetter,[5] Lucy S. Petro,[1,2,6] and Lars F. Muckli[1,2,7,*]

[1]Centre for Cognitive Neuroimaging, School of Psychology and Neuroscience, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QB, UK
[2]Imaging Centre for Excellence (ICE), College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G51 4LB, UK
[3]School of Computing Science, College of Science and Engineering, University of Glasgow, Glasgow G12 8QQ, UK
[4]Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1, Leipzig 04303, Germany
[5]Visual & Cognitive Neuroscience Lab, Department of Psychology, University of Fribourg, Fribourg 1700, Switzerland
[6]These authors contributed equally
[7]Lead contact
*Correspondence: lars.muckli@glasgow.ac.uk
https://doi.org/10.1016/j.cub.2024.09.008

SUMMARY

Listening to natural auditory scenes leads to distinct neuronal activity patterns in the early visual cortex (EVC) of blindfolded sighted and congenitally blind participants.[1,2] This pattern of sound decoding is organized by eccentricity, with the accuracy of auditory information increasing from foveal to far peripheral retinotopic regions in the EVC (V1, V2, and V3). This functional organization by eccentricity is predicted by primate anatomical connectivity,[3,4] where cortical feedback projections from auditory and other non-visual areas preferentially target the periphery of early visual areas. In congenitally blind participants, top-down feedback projections to the visual cortex proliferate,[5] which might account for even higher sound-decoding accuracy in the EVC compared with blindfolded sighted participants.[2] In contrast, studies in participants with aphantasia suggest an impairment of feedback projections to early visual areas, leading to a loss of visual imagery experience.[6,7] This raises the question of whether impaired visual feedback pathways in aphantasia also reduce the transmission of auditory information to early visual areas. We presented auditory scenes to 23 blindfolded aphantasic participants. We found overall decreased sound decoding in early visual areas compared to blindfolded sighted ("control") and blind participants. We further explored this difference by modeling eccentricity effects across the blindfolded control, blind, and aphantasia datasets, and with a whole-brain searchlight analysis. Our findings suggest that the feedback of auditory content to the EVC is reduced in aphantasic participants. Reduced top-down projections might lead to both less sound decoding and reduced subjective experience of visual imagery.

RESULTS

Sound decoding in the early visual cortex of aphantasic participants

We first confirmed our participants' self-reported aphantasia. Twenty-three participants scored low on the vividness of visual imagery questionnaire (VVIQ[8,9]) (mean = 20.40, SD = 4.29), indicating that they were aphantasic.[10–12] Participants' spontaneous use of imagery scale (SUIS[13]) scores were also low (mean = 19.13, SD = 4.07), in line with previous reports.[14] We additionally performed a behavioral pre-test confirming that imagery priming in aphantasic participants differed from typical imagers[15] (STAR Methods; Figure S1).

Next, during 3T fMRI, we presented blindfolded aphantasic participants with 3 auditory scenes (forest, crowd, and traffic), replicating our previous design in blindfolded sighted participants[1] (Figure 1, STAR Methods). In line with our previous findings,[1,2] univariate blood-oxygen-level-dependent (BOLD) responses were weakly deactivated in our participants' retinotopic areas (early visual cortex [EVC], V1, V2, and V3), in the areas' eccentricity fields (fovea, periphery, and far periphery), and in the motor cortex, but were strongly activated in the auditory cortex (see Figure S2).

Using multivariate pattern analyses, we tested for sound decoding from activity patterns in V1, V2, and V3. We used auditory and motor areas as positive and negative controls, respectively. We could not decode sound content significantly in any early visual area when collapsing across eccentricity (EVC = 34.06%, $p > 0.05$; V1 = 35.14%, $p > 0.05$; V2 = 32.61%, $p > 0.05$; V3 = 33.70%, $p > 0.05$; chance level = 33%; Figure 2A). When delineating areas V1, V2, and V3 by eccentricity (fovea, periphery, and far periphery), we found non-significant decoding accuracies in most regions (EVC fovea = 32.97%, EVC periphery = 30.07%, EVC far periphery = 33.33%, V1 fovea = 28.99%, V1 periphery = 33.70%, V2 fovea = 31.16%, V2 periphery = 30.43%, V2 far periphery = 31.88%, V3 periphery = 29.71%, V3 far periphery = 32.61%; Figure 2B), except far peripheral V1 (38.41%, $p < 0.05$) and foveal V3 (39.13%, $p < 0.05$; see Figure 2B). This

Please cite this article in press as: Montabes de la Cruz et al., Decoding sound content in the early visual cortex of aphantasic participants, Current Biology (2024), https://doi.org/10.1016/j.cub.2024.09.008

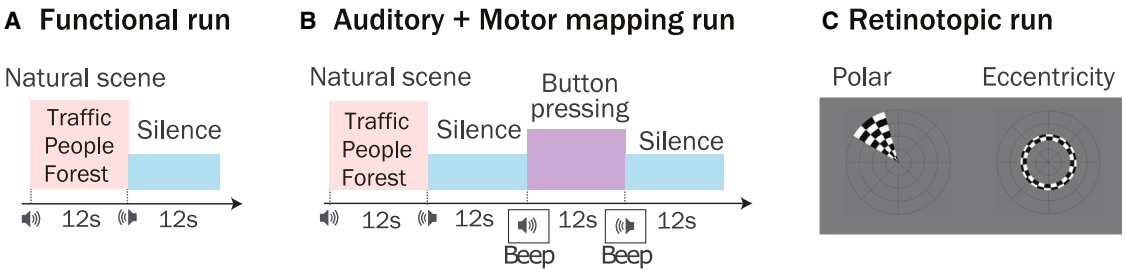

Figure 1. Experimental methods

(A) In our functional runs, participants were presented with a 12-s-long natural sound scene (traffic scene or non-semantic people talking scene or forest scene), which was followed by 12 s of silence. Each sound was presented 6 times in pseudo-randomized order. No scene was repeated immediately after itself.

(B) In our auditory and motor-mapping run, the paradigm was identical, except that we replaced half of the blocks with motor-area-localizing blocks. During these, participants were instructed to press the buttons of a response box in ascending order. In both runs, participants were blindfolded and the lights in the scanner room were switched off.

(C) Before commencing the retinotopic run, participants took off their blindfold and were presented with the polar and eccentricity stimuli. As an attentional control, participants were instructed to press a button whenever a central dot changed color. See Figure S1 for behavioral methods and results.

suggests that sound decoding in aphantasic participants is attenuated compared with blindfolded sighted and blind participants, where in both cases it was stronger and more widespread in early visual areas (Figures 2C–2F). In control regions, decoding was successful in the auditory cortex (98.91%, $p < 0.001$) but not in the motor cortex (36.59%, $p > 0.05$; Figure 2A), consistent with our previous control study.[1] Finally, we calculated Bayes factors and interpreted them using Lee and Wagenmakers' scale.[16] The results were consistent with our permutation test p values, showing moderate support for H1 (i.e., sound decoding) in V1 far periphery and V3 fovea, and at least moderate support for H0 (i.e., no sound decoding) in the remaining EVC areas (see Table S1).

The gradient of sound decoding accuracy and eccentricity might reflect the strength of feedback connectivity to EVC

We further explored sound decoding profiles by modeling eccentricity effects across the control,[1] blind[2] and aphantasia participants. We fitted a linear mixed model to predict classification accuracy scores by group, area, and eccentricity. When averaging across areas and eccentricities, we found a significant variation in accuracy between groups (Figure 3A.1; $F(2, 38) = 13.45$; $p < 0.0001$), with the blind group showing the highest accuracy and the aphantasic group the lowest. When averaging across groups and areas, we found a significant accuracy trend by eccentricity (Figure 3A.2; $F(1, 38) = 6.07$; $p < 0.05$), with far peripheral fields showing higher decoding than foveal fields. However, when averaging across eccentricities and groups, the main effect of retinotopic areas ($F(2, 275) = 0.49$; $p > 0.1$) was non-significant, indicating no overall difference in accuracy between V1, V2, and V3. Additionally, when averaging across areas, we found a significant interaction between group and eccentricity (Figure 3A.3; $F(2, 38) = 3.82$; $p < 0.05$), where a trend of higher accuracy in the far periphery was the most marked for the blind group and least for the aphantasia group.

When averaging across groups, we found a significant interaction between area and eccentricity (Figure 3A.4; $F(2, 275) = 4.88$; $p < 0.01$), where the positive eccentricity trend was the most marked for V1 and least for V3. The interaction effect between group and area was non-significant ($F(4, 275) = 0.83$; $p > 0.1$), indicating that, when averaging over eccentricities and controlling for the main effect of group, there was no strong group-wise pattern in the variations of accuracy between V1, V2, and V3. Finally, we found that the interaction between group, eccentricity, and area was also non-significant ($F(4, 275) = 0.35$; $p > 0.1$), suggesting that first-level interactions were sufficient to capture the complexity of our data.

To further explore how each group and region contributed to these effects, we derived estimated marginal means for each combination of predictors, which we used for pairwise comparisons of accuracy and eccentricity trends (Figure 3A). Across areas, we found that sound decoding accuracy significantly differed between the blind and the two other groups (Figure 3A.1; aphantasia–blind: $t(38) = -5.16$; $p < .0001$, aphantasia–control: $t(38) = -1.98$; $p > 0.1$, blind–control: $t(38) = 2.88$, $p < 0.05$), while the eccentricity trend only significantly differed between aphantasic and blind participants (Figure 3A.3; aphantasia–blind: $t(38) = -2.76$; $p < 0.05$, aphantasia–control: $t(38) = -0.93$; $p > 0.1$, blind– control: $t(38) = 1.65$; $p > 0.1$). Across groups, we found no significant difference between areas in terms of accuracy (Figure 3A.4; V1–V2: $t(275) = 0.58$; $p > 0.1$, V1–V3: $t(275) = 1.26$; $p > 0.1$, V2–V3: $t(275) = 0.68$; $p > 0.1$), but there was a significant difference between V3 and V1, and a tendential difference between V3 and V2, in eccentricity trend (V1–V2: $t(275) = 0.37$; $p > 0.1$, V1–V3: $t(275) = 2.56$; $p < 0.05$, V2–V3: $t(275) = 2.19$; $p < 0.1$).

This model analysis suggests that auditory scenes are more strongly represented in early visual areas (both in terms of decoding accuracy and in the increase of accuracy across eccentricities) when top-down feedback projections to the EVC proliferate (e.g., in congenitally blind participants) than when these projections may be impaired (e.g., in aphantasia). Moreover, increased sound decoding with eccentricity is most prominent in V1 compared with V3.

Reduced feedback connectivity in aphantasia may be EVC specific

We performed a searchlight analysis using our 3 datasets (aphantasia, blind, and sighted) to test for auditory scene

Please cite this article in press as: Montabes de la Cruz et al., Decoding sound content in the early visual cortex of aphantasic participants, Current Biology (2024), https://doi.org/10.1016/j.cub.2024.09.008

CellPress
OPEN ACCESS

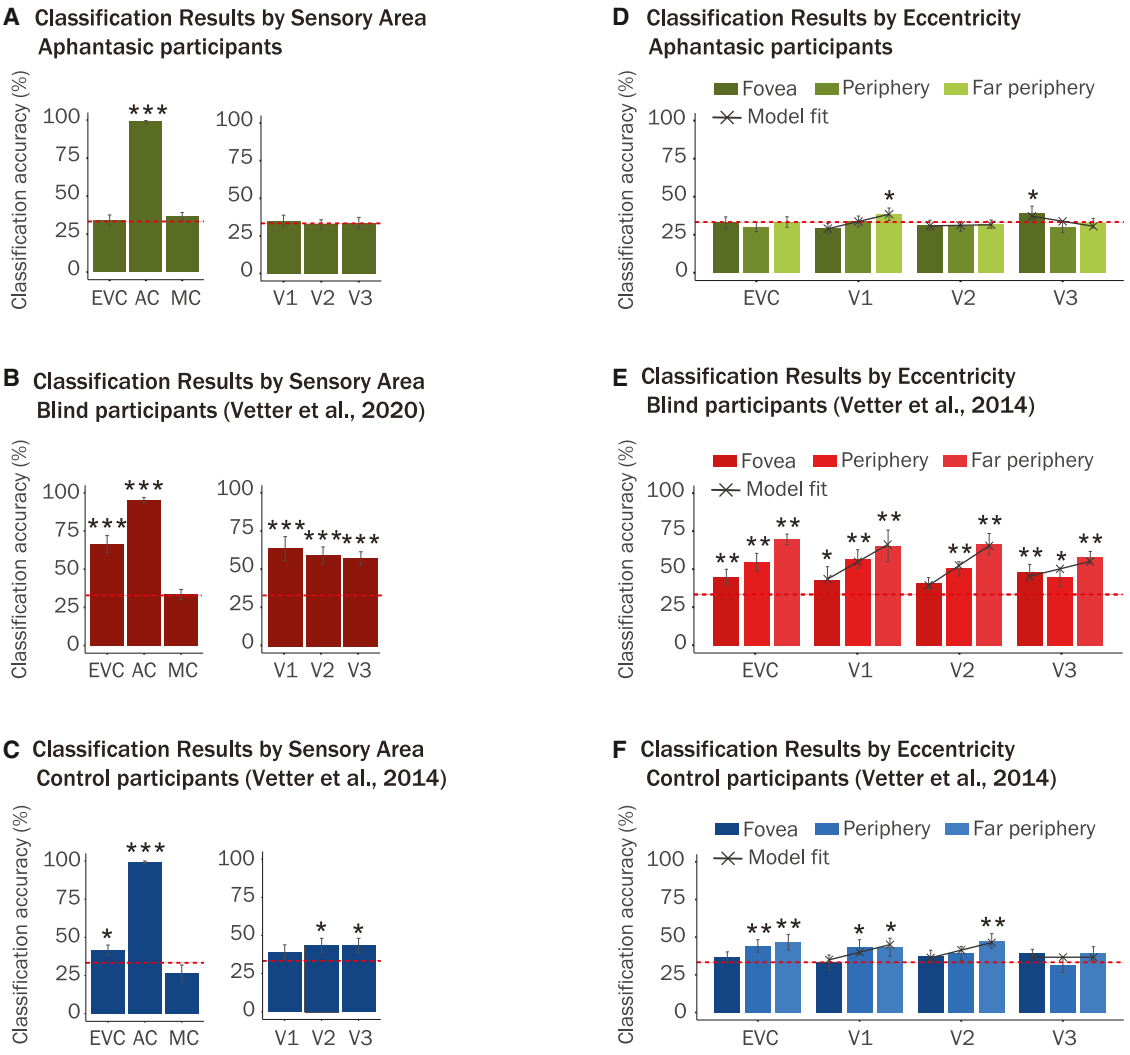

Figure 2. Multi-voxel pattern analysis results

Average classification accuracies for aphantasics (A and D), blind participants (B and E), and controls (C and F). (A)–(C): whole-ROI decoding in auditory and motor cortex as well as V1, V2, and V3. (D)–(F): average classification accuracies by eccentricity (fovea, periphery, and far periphery) in the averaged early visual cortex, V1, V2, and V3. Horizontal dashed line indicates chance level (33%) and error bars indicate SEM. p values derived from our permutation analyses. Significance threshold was established as $p < 0.05$, where ** indicates $p < 0.01$ and *** indicates $p < 0.001$. See Figure S2 for univariate results and Table S1 for the Bayes factors of our reported classification accuracies.

decoding beyond the EVC (Figures 3B and 3C). As expected, our averaged searchlight maps showed the highest level of decoding in the auditory cortex and in multisensory areas in all groups. Averaged searchlight maps did not differ between participants with aphantasia and controls (Figures 3B and 3C, decoding was similarly high and clustered in auditory and multisensory areas) but were different to blind participants (see next paragraph). Region of interest (ROI)-based multi-voxel pattern analysis results suggested a reduced representation of sounds in the EVC in participants with aphantasia (we only observed decoding in peripheral V1 and foveal V3). In controls[1] and, especially, in the blind group,[2] we reported a more widespread pattern of decoding in the EVC. Decreased sound decoding in aphantasia could indicate that top-down feedback projections from auditory cortex to EVC are reduced. These projections may reach peripheral V1 without carrying over to peripheral V2, unlike in controls.[1] Moreover, the fact that our searchlight analysis does not show differences in sound-decoding elsewhere suggests that the presumed differences in connectivity do not extend to other brain areas. If there is reduced feedback connectivity in aphantasia, it may be specific to top-down feedback to the EVC, as suggested by models of visual aphantasia.[17]

By contrast, in blind participants (Figure 3D), decoding in the auditory and multisensory cortex was lower relative to the other two groups (Figures 3B and 3C). We previously reported this finding of lower sound decoding in auditory areas in the blind group.[2] We now extend our findings by also reporting lower sound decoding in multisensory areas in the blind group and by replicating the comparison with aphantasic participants (i.e., comparable sound decoding between controls and

Please cite this article in press as: Montabes de la Cruz et al., Decoding sound content in the early visual cortex of aphantasic participants, Current Biology (2024), https://doi.org/10.1016/j.cub.2024.09.008

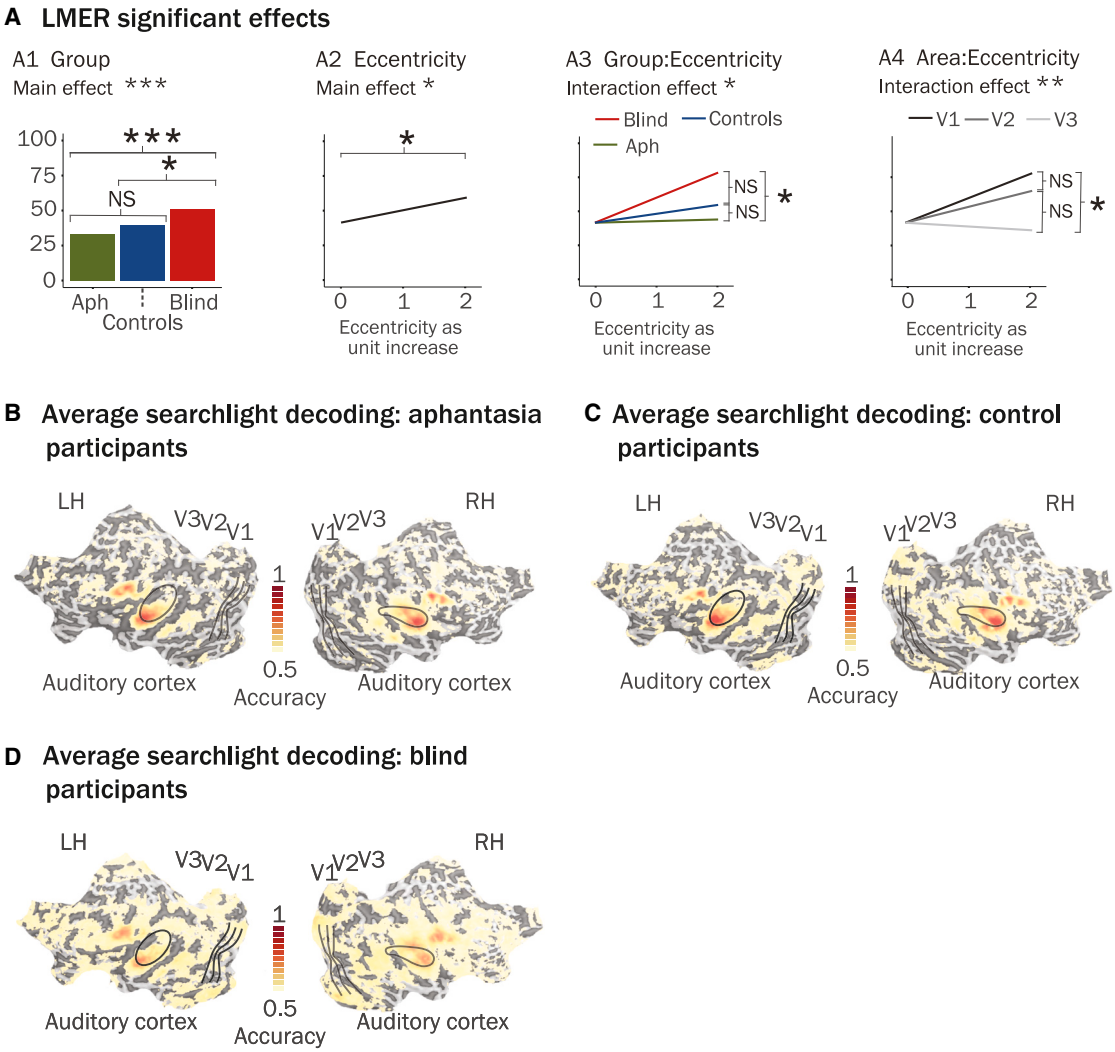

Figure 3. Group comparisons and searchlight results

(A) Significant main effects reported by an ANOVA analysis of our LMER model, which included (A1) a main effect of group, (A2) a main effect of eccentricity, (A3) an interaction effect between group and eccentricity, and (A4) an interaction effect between area and eccentricity. All figures include pairwise comparisons. Significance threshold * was established as $p < 0.05$, where ** indicates $p < 0.01$ and *** indicates $p < 0.001$.

(B–D) Averaged searchlight decoding for our (B) aphantasia group, (C) Vetter et al.'s[1] control group, and (D) Vetter et al.'s[2] congenitally blind group. Light yellow coloring indicates chance performance (0.5) and dark red coloring indicates perfect performance (1).

aphantasic groups). Sound decoding, as indicated by average searchlight maps, is more widespread in the blind group than in the control and the aphantasic groups. In the blind group, searchlight decoding included visual regions, while in controls and aphantasic participants, decoding was stronger in auditory and multisensory areas. This pattern of results in the blind group suggests a shift in the distribution of sound representation from auditory to include visual cortices.

DISCUSSION

We have previously shown that listening to auditory scenes leads to neuronal activity patterns in the EVC of blindfolded sighted and congenitally blind participants. Moreover, this pattern is dependent on the retinotopic organization of early visual areas, with auditory sound decoding increasing with eccentricity.[1,2] Due to cross-modal plasticity and proliferation of feedback projections, non-feedforward-driven activity in the visual cortex of congenitally blind participants is greater compared with sighted participants.[5] Here, we tested aphantasic participants, a group suggested to have impaired visual feedback pathways to the EVC.[14,17] If, in accordance with our previous results, sound decoding in the EVC is independent of visual imagery, then it should be preserved in aphantasia. However, given that we found increased decoding for congenitally blind participants compared with sighted controls, the relative strength of feedback projections to the EVC might be what determines the ability to decode sounds in the EVC. If this were the case, sound decoding in the EVC would be generally attenuated in participants with impaired feedback projections.

Please cite this article in press as: Montabes de la Cruz et al., Decoding sound content in the early visual cortex of aphantasic participants, Current Biology (2024), https://doi.org/10.1016/j.cub.2024.09.008

In our decoding analyses in aphantasic participants, we find evidence of sound decoding in peripheral V1, suggesting that feedback of auditory information is independent of imagery. However, we also find that sound decoding in aphantasic participants is attenuated compared with blindfolded sighted and blind participants. Imagery content can be decoded from early visual areas in typical imagers,[18,19] questioning whether the reduction of feedback connectivity in aphantasia is specific to an imagery-related signal. In a control experiment on Vetter et al.'s[1] blindfolded sighted participants, imagined auditory scenes (as opposed to passive listening) could be decoded in the foveal and peripheral area of the EVC, but not in the far periphery.[1] This pattern of decoding echoes other findings of foveal decoding for some types of task-relevant feedback.[20,21] In contrast, more peripheral decoding during passive listening[1] (see Figure 2F) aligns with reports showing that auditory attention suppresses foveal and enhances peripheral early visual representations.[22] We could therefore expect that decreasing imagery-related feedback would increase the auditory decoding with an eccentricity trend from the fovea to periphery in the aphantasia group. Instead, we found that the eccentricity trend, much like decoding accuracy in general, increased from aphantasic to control to blind groups, being significantly different between blind and aphantasic participants. This would rather indicate a reduced multimodal feedback from the auditory to the visual cortex, which we know preferentially targets the periphery.[3,4] Finally, recent evidence suggests that the EVC is not generally engaged during imagery of high-level semantic content, such as in our auditory scenes.[23] Another result emerging from our model is that the eccentricity trend progressively decreases when looking at higher areas in the visual hierarchy, being significantly lower in V3 compared with V1. This could be a consequence of higher areas becoming less retinotopic as they specialize in processing other types of information.

Current research on the neural bases of aphantasia debates whether aphantasia is primarily a visual consciousness disorder making individuals unaware of their imagery representations,[7] whether impaired feedback prevents aphantasic participants from generating representations,[14] or whether aphantasia is due to an episodic memory impairment.[24–26] Other evidence suggests that there may be more nuance behind aphantasia. Meng et al.[27] reported similar visual imagery decoding in V1 between aphantasic and control groups but different cross-decoding with perception, suggesting that, if present, aphantasic representations may differ from those of typical imagers. By contrast, Cabbai et al.[28] reported imagery decoding in V1 in controls but not in aphantasia participants. Additionally, in a similar design to ours, the authors reported comparable V1 decoding between aphantasic and control groups in response to natural sounds.[28] In our case, we do not find decoding in response to sounds using V1 as a whole, but instead find that the sounds were represented in the areas' periphery only, which, although not a direct measure of anatomical connectivity, contributes to the argument in favor of reduced functional feedback processing.

Overall, our results suggest that aphantasia might be associated with decreased feedback connectivity that is specific to the EVC, some of which at least is not related to imagery but to cross-modal feedback signals. By contrast, the searchlight maps of congenitally blind participants show lower decoding in auditory and multisensory areas and more widespread decoding in visual areas. This observation indicates more feedback reorganization in the blind participants, where auditory processing may be shifted to visual areas to compensate for the group's absent visual feedforward processing.[2] As such, our results support the notion that the three groups might differ in feedback connectivity to the early visual areas. The blind group exhibits the strongest connectivity profile, the aphantasic group the weakest, with controls being in between. These findings suggest that aphantasic participants may respond differently to controls in other auditory-driven processes. For instance, aphantasic participants could be less susceptible to audio-visual illusions (e.g., contingent auditory motion aftereffect[29]). Moreover, aphantasic participants could be expected to exhibit less-pronounced cross-modal effects during sound localization.[30] The retinotopic structure of the early visual areas is suited to process spatial properties,[31] such as during the localization of sound sources (in our case, forest/birds sounds above, human sounds at eye-level and in the vicinity, and car sounds from the ground and in the distance). However, if the communication from auditory to early visual areas is impaired, as our findings and those of others suggest to be the case in aphantasia, we could expect sound localization responses to differ. This could be expressed as slower, but not less-accurate responses, as seen in a previous study,[32] and as a reduced recruitment of the EVC during sound localization in aphantasia. The link between decodability of natural sounds and space cues and its relatedness to aphantasia needs to be investigated in more targeted follow-up studies.

RESOURCE AVAILABILITY

Lead contact

Requests for resources or further information should be directed to the lead contact, Lars Muckli (Lars.Muckli@glasgow.ac.uk).

Materials availability

This study did not generate new materials.

Data and code availability

The code to run the experiment is available on https://github.com/Muckli-lab/NaturalSound-Stimulation and that to perform the multi-voxel pattern analysis is available on https://github.com/Muckli-lab/MVP-analysis-tool-box.

The current MRI dataset is available on the OpenNeuro platform under https://openneuro.org/datasets/ds005562/versions/1.0.0. The MRI dataset from Vetter et al.[2] is accessible on the EBRAINS platform of the Human Brain Project under https://kg.ebrains.eu/search/instances/Dataset/4f6e1509-2e7f-44dd-a45c-c100cd7728a3.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (awarded to L.F.M.; ERC-StG-2012-311751 BrainReadFBPredCode); the European Union's Horizon 2020 Framework Programme for Research and Innovation under the specific grant agreement nos. 720270, 785907, and 945539 (Human Brain Project SGA1, SGA2, and SGA3); and the Biotechnology and Biological Sciences Research Council (BBSRC, BBN010956/1) "Layer-specific cortical feedback" to L.F.M. and L.S.P., and was also funded by a PRIMA grant (PR00P1_185918/1) from the Swiss National Science Foundation to P.V. and through the College of Medicine, Veterinary and Life Sciences doctoral training programme (MVLS DTP) at the University of Glasgow for B.M.M.d.l.C. We thank Frances Crabbe, Zirui Zang, and Bohdana Neurathova for help in

acquiring fMRI data and Adam Zeman for help with participant recruitment. We thank Amir Amedi and Lior Reich for their contribution to the dataset from congenitally blind participants and all authors listed on Vetter et al.[1,2] For the purpose of open access, the author(s) has applied a Creative Commons Attribution (CC BY) license to any author accepted manuscript version arising from this submission.

AUTHOR CONTRIBUTIONS

L.F.M., B.M.M.d.l.C., C.A., L.S.P., and P.V. conceived the study; B.M.M.d.l.C., C.A., and R.S.L. collected the data; B.M.M.d.l.C., C.A., R.S.L., and A.T.P. analyzed the data; P.V., J.B., and L.F.M. provided comparison datasets; B.M.M.d.l.C., C.A., L.S.P., R.S.L., A.T.P., J.B., and L.F.M. wrote and revised the manuscript; L.F.M. and L.S.P acquired funding.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Questionnaires to screen for aphantasia
  - Behavioural testing
  - fMRI
  - Data acquisition
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Imagery questionnaires, behavioural testing, and imagery priming results
  - Pre-processing, cortical surface reconstruction, and region-of-interest definition
  - Univariate analysis
  - Multivariate pattern classification analysis
  - Bayes factors of classification accuracies
  - Eccentricity modelling in EVC across populations
  - Whole brain searchlight analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2024.09.008.

Received: July 1, 2024
Revised: August 21, 2024
Accepted: September 4, 2024
Published: October 16, 2024

REFERENCES

1. Vetter, P., Smith, F.W., and Muckli, L. (2014). Decoding Sound and Imagery Content in Early Visual Cortex. Curr. Biol. *24*, 1256–1262. https://doi.org/10.1016/j.cub.2014.04.020.
2. Vetter, P., Bola, Ł., Reich, L., Bennett, M., Muckli, L., and Amedi, A. (2020). Decoding Natural Sounds in Early "Visual" Cortex of Congenitally Blind Individuals. Curr. Biol. *30*, 3039–3044.e2. https://doi.org/10.1016/j.cub.2020.05.071.
3. Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. J. Neurosci. *22*, 5749–5759. https://doi.org/10.1523/JNEUROSCI.22-13-05749.2002.
4. Wang, M., Hou, Y., Magrou, L., Autio, J.A., Misery, P., Coalson, T., Reid, E., Xu, Y., Lamy, C., Falchier, A., et al. (2022). Retinotopic organization of feedback projections in primate early visual cortex: implications for active vision. Preprint at bioRxiv. https://doi.org/10.1101/2022.04.27.489651.
5. Bedny, M. (2017). Evidence from Blindness for a Cognitively Pluripotent Cortex. Trends Cogn. Sci. *21*, 637–648. https://doi.org/10.1016/j.tics.2017.06.003.
6. Milton, F., Fulford, J., Dance, C., Gaddum, J., Heuerman-Williamson, B., Jones, K., Knight, K.F., MacKisack, M., Winlove, C., and Zeman, A. (2021). Behavioral and Neural Signatures of Visual Imagery Vividness Extremes: Aphantasia versus Hyperphantasia. Cereb. Cortex Commun. *2*, tgab035. https://doi.org/10.1093/texcom/tgab035.
7. Liu, J., Zhan, M., Hajhajate, D., Spagna, A., Dehaene, S., Cohen, L., and Bartolomeo, P. (2023). Visual mental imagery in typical imagers and in aphantasia: A millimeter-scale 7-T fMRI study. Preprint at bioRxiv. https://doi.org/10.1101/2023.06.14.544909.
8. Marks, D.F. (1973). Vividness of visual imagery Questionnaire. J. Ment. Imagery. https://doi.org/10.1037/t05959-000.
9. Marks, D.F. (1995). New directions for mental imagery research. J. Ment. Imagery *19*, 153–167.
10. Zeman, A., Dewar, M., and Della Sala, S. (2015). Lives without imagery – Congenital aphantasia. Cortex *73*, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019.
11. Dance, C.J., Ipser, A., and Simner, J. (2022). The prevalence of aphantasia (imagery weakness) in the general population. Conscious. Cogn. *97*, 103243. https://doi.org/10.1016/j.concog.2021.103243.
12. Takahashi, J., Saito, G., Omura, K., Yasunaga, D., Sugimura, S., Sakamoto, S., Horikawa, T., and Gyoba, J. (2023). Diversity of aphantasia revealed by multiple assessments of visual imagery, multisensory imagery, and cognitive style. Front. Psychol. *14*, 1174873. https://doi.org/10.3389/fpsyg.2023.1174873.
13. Kosslyn, S.M., Chabris, C., Shephard, J., and Thompson, W. (2017). Spontaneous Use of Imagery Scale. Applied Cognitive Psychology. https://doi.org/10.1037/t57899-000.
14. Keogh, R., and Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. Cortex J. Devoted Study Nerv. Syst. Behav. *105*, 53–60. https://doi.org/10.1016/j.cortex.2017.10.012.
15. Bergmann, J., Petro, L.S., Abbatecola, C., Li, M.S., Morgan, A.T., and Muckli, L. (2024). Cortical depth profiles in primary visual cortex for illusory and imaginary experiences. Nat. Commun. *15*, 1002. https://doi.org/10.1038/s41467-024-45065-w.
16. Lee, M.D., and Wagenmakers, E.-J. (2013). Bayesian Cognitive Modeling: A Practical Course (Cambridge University Press).
17. Dawes, A.J., Keogh, R., and Pearson, J. (2024). Multisensory subtypes of aphantasia: Mental imagery as supramodal perception in reverse. Neurosci. Res. *201*, 50–59. https://doi.org/10.1016/j.neures.2023.11.009.
18. Ragni, F., Tucciarelli, R., Andersson, P., and Lingnau, A. (2020). Decoding stimulus identity in occipital, parietal and inferotemporal cortices during visual mental imagery. Cortex *127*, 371–387. https://doi.org/10.1016/j.cortex.2020.02.020.
19. Ragni, F., Lingnau, A., and Turella, L. (2021). Decoding category and familiarity information during visual imagery. Neuroimage *241*, 118428. https://doi.org/10.1016/j.neuroimage.2021.118428.
20. Williams, M.A., Baker, C.I., Op de Beeck, H.P., Shim, W.M., Dang, S., Triantafyllou, C., and Kanwisher, N. (2008). Feedback of visual object information to foveal retinotopic cortex. Nat. Neurosci. *11*, 1439–1445. https://doi.org/10.1038/nn.2218.
21. Bennett, M.A., Petro, L.S., Abbatecola, C., and Muckli, L. (2024). Retinotopic biases in contextual feedback signals to V1 for object and scene processing. Preprint at bioRxiv. https://doi.org/10.1101/2024.03.26.586553.

22. Gau, R., Bazin, P.L., Trampel, R., Turner, R., and Noppeney, U. (2020). Resolving multisensory and attentional influences across cortical depth in sensory cortices. eLife *9*, e46856. https://doi.org/10.7554/eLife.46856.
23. Dijkstra, N. (2024). Uncovering the Role of the Early Visual Cortex in Visual Mental Imagery. Vision (Basel) *8*, 29. https://doi.org/10.3390/vision8020029.
24. Blomkvist, A. (2023). Aphantasia: In search of a theory. Mind Lang. *38*, 866–888. https://doi.org/10.1111/mila.12432.
25. Monzel, M., Leelaarporn, P., Lutz, T., Schultz, J., Brunheim, S., Reuter, M., and McCormick, C. (2023). Hippocampal-Occipital Connectivity Reflects Autobiographical Memory Deficits in Aphantasia (Cold Spring Harbor Laboratory).
26. Bergmann, J., and Ortiz-Tudela, J. (2023). Feedback signals in visual cortex during episodic and schematic memory retrieval and their potential implications for aphantasia. Neurosci. Biobehav. Rev. *152*, 105335. https://doi.org/10.1016/j.neubiorev.2023.105335.
27. Meng, M., Chang, S., Zhang, X., and Pearson, J. (2023). Imageless imagery in aphantasia: decoding non-sensory imagery in aphantasia. Res. Square. https://doi.org/10.21203/rs.3.rs-3162223/v1.
28. Cabbai, G., Racey, C., Simner, J., Dance, C., Ward, J., and Forster, S. (2024). Sensory representations in primary visual cortex are not sufficient for subjective imagery. Preprint at bioRxiv. https://doi.org/10.1101/2024.01.10.574972.
29. Dong, C., Swindale, N.V., and Cynader, M.S. (1999). A contingent aftereffect in the auditory system. Nat. Neurosci. *2*, 863–865. https://doi.org/10.1038/13161.
30. van der Heijden, K., Rauschecker, J.P., Formisano, E., Valente, G., and de Gelder, B. (2018). Active Sound Localization Sharpens Spatial Tuning in Human Primary Auditory Cortex. J. Neurosci. *38*, 8574–8587. https://doi.org/10.1523/JNEUROSCI.0587-18.2018.
31. Haun, A., and Tononi, G. (2019). Why Does Space Feel the Way it Does? Towards a Principled Account of Spatial Experience. Entropy *21*, https://doi.org/10.3390/e21121160.
32. Kay, L., Keogh, R., and Pearson, J. (2024). Slower but more accurate mental rotation performance in aphantasia linked to differences in cognitive strategies. Conscious. Cogn. *121*, 103694. https://doi.org/10.1016/j.concog.2024.103694.
33. Goebel, R., Esposito, F., and Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. Hum. Brain Mapp. *27*, 392–401. https://doi.org/10.1002/hbm.20249.
34. Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. *2*, 1–27. https://doi.org/10.1145/1961189.1961199.
35. Team, R.C. (2018). R: a language and environment for statistical computing R Foundation for Statistical Computing. Computing.
36. Team, R.S. (2020). RStudio: Integrated Development Environment for R (RStudio, PBC).
37. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Software *67*, 1–48. https://doi.org/10.18637/jss.v067.i01.
38. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. J. Stat. Software *82*, 1–26. https://doi.org/10.18637/jss.v082.i13.
39. Lenth, R.V. (2017). Emmeans: Estimated Marginal Means, aka Least-Squares Means. https://doi.org/10.32614/CRAN.package.emmeans.
40. Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in psychtoolbox-3. Perception *36*, 1.
41. Pearson, J., Clifford, C.W.G., and Tong, F. (2008). The functional impact of mental imagery on conscious perception. Curr. Biol. *18*, 982–986. https://doi.org/10.1016/j.cub.2008.05.048.
42. Wandell, B.A., Dumoulin, S.O., and Brewer, A.A. (2007). Visual field maps in human cortex. Neuron *56*, 366–383. https://doi.org/10.1016/j.neuron.2007.10.012.
43. Schira, M.M., Tyler, C.W., Breakspear, M., and Spehar, B. (2009). The foveal confluence in human visual cortex. J. Neurosci. *29*, 9050–9058. https://doi.org/10.1523/JNEUROSCI.1760-09.2009.
44. Keogh, R., Wicken, M., and Pearson, J. (2021). Visual working memory in aphantasia: Retained accuracy and capacity with a different strategy. Cortex *143*, 237–253. https://doi.org/10.1016/j.cortex.2021.07.012.
45. Kay, L., Keogh, R., Andrillon, T., and Pearson, J. (2022). The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength. eLife *11*, e72484. https://doi.org/10.7554/eLife.72484.
46. Lorena, A.C., de Carvalho, A.C.P.L.F., and Gama, J.M.P. (2008). A review on the combination of binary classifiers in multiclass problems. Artif. Intell. Rev. *30*, 19–37. https://doi.org/10.1007/s10462-009-9114-9.

Please cite this article in press as: Montabes de la Cruz et al., Decoding sound content in the early visual cortex of aphantasic participants, Current Biology (2024), https://doi.org/10.1016/j.cub.2024.09.008

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
MRI data	OpenNeuro.org, Ebrains.eu	Montabes de la Cruz et al. https://openneuro.org/datasets/ds005562/versions/1.0.0; Vetter et al.[2] https://search.kg.ebrains.eu/instances/Dataset/4f6e1509-2e7f-44dd-a45c-c100cd7728a3
Experimental code andsound data files	Github	https://github.com/Muckli-lab/NaturalSound-Stimulation
MVPA analysis code	Github	https://github.com/Muckli-lab/MVP-analysis-tool-box
Software and algorithms		
Presentation (23.0)	Neurobehavioral Systems	https://www.neurobs.com; RRID: SCR_002521
Brainvoyager (22.4)	BrainInnovation, Maastricht[33]	http://www.brainvoyager.com; RRID: SCR_013057
Matlab (R2016a, R2019a)	MathWorks	https://www.mathworks.com; RRID: SCR_001622
NeuroElf (v11)	Neuroelf.net	https://neuroelf.net/; RRID: SCR_014147
LIBSVM (3.3)	Chang et al.[34]	https://www.csie.ntu.edu.tw/~cjlin/libsvm/; RRID: SCR_010243
R (4.3 - 4.4)	R Core Team[35]	https://www.R-project.org/; RRID: SCR_001905
RStudio (2024.04.2)	RStudio Team[36]	https://posit.co/; RRID: SCR_000432
lme4	Bates et al.[37]	https://cran.r-project.org/web/packages/lme4/index.html; RRID: SCR_015654
lmerTest	Kuznetsova et al.[38]	https://cran.r-project.org/web/packages/lmerTest/; RRID: SCR_015656
emmeans	Lenth[39]	https://cran.r-project.org/web/packages/emmeans/index.html; RRID: SCR_018734
Psychtoolbox v3.0.13	Kleiner et al.[40]	http://psychtoolbox.org/; RRID: SCR_002881

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

23 (from 31 participants, 8 of which did not pass fMRI screening or withdrew from the study) self-reported aphantasic participants gave informed consent to participate in the experiment (16 female, 6 male, 1 non-disclosed, mean age = 30.9 years, range=19-64 years). The College of Science and Engineering of the University of Glasgow granted ethical approval (application number 300190011). Participants received monetary compensation. The behavioural experiment lasted for one hour and the fMRI session lasted one and a half hours. We conducted behavioural testing and brain imaging on separate days, at the Center of Cognitive Neuroimaging (CCNi) of the University of Glasgow.

We used the datasets of 10 participants (7 females, mean age= 24.1 years, range = 20-33 years) from Experiment 1 of Vetter et al.[1] as the control group, and of 8 participants (5 females, mean age = 33.4 years, range = 23-39 years) from Vetter et al.[2] as the congenitally blind group.

METHOD DETAILS

Questionnaires to screen for aphantasia

Participants completed the vividness of visual imagery questionnaire (VVIQ),[8,9] which instructs participants to imagine four scenes (eg. rising sun) and then rank how vividly they perceive four additional aspects of the scenes (eg. "The sun is rising above the horizon into a hazy sky"). Rankings are scaled from 1 ("No image at all, you only "know" that you are thinking of the object") to 5 ("Perfectly clear and as vivid as normal vision"). Additionally, participants completed the spontaneous use of imagery scale (SUIS),[13] which instructs participants to state how appropriate certain statements are in relation to their experience of more automatic forms of visual imagery (eg. "When I think about a series of errands I must do, I visualise the stores I will visit"). It is ranked from 1 (never appropriate) to 5 (always appropriate).

Please cite this article in press as: Montabes de la Cruz et al., Decoding sound content in the early visual cortex of aphantasic participants, Current Biology (2024), https://doi.org/10.1016/j.cub.2024.09.008

Behavioural testing

The imagery priming test assesses the extent to which two imagined stimuli can prime their subsequent rivalrous perception. Stimuli and procedures of the binocular rivalry paradigm were the same as in Keogh and Pearson's[14] no luminance condition in all participants but 5, for whom we delivered monocular stimulation using shutter glasses instead of coloured glasses as described in their study.[14] We controlled the visual stimulation using MATLAB R2016a in combination with the Psychtoolbox v3.0.13 extension.[40] We fixed participants' viewing distance from the monitor screen at 60cm with a chin rest and they wore red/green glasses to ensure monocular stimulation. We displayed stimuli on a black background around a fixation point. Stimuli consisted of a red horizontal (CIE x =.57, y =.36) and green vertical (CIE x =.28, y =.63) Gabor patterns (1 cycle/°, Gaussian $\sigma = 1.5°$, mean luminance 4.41 cdm2) presented in an annulus, as well as mock rivalry displays (12.5% of trials, half red and half green) to assess any decisional bias in the imagery task. Participants first completed an eye dominance task to prevent individual differences in eye dominance from biasing the rivalrous perception,[41] then a training run (10 trials) to become familiar with the task, following which they engaged in the final imagery run (100 trials with a break at 50). During the trial, a grey letter 'R' (red) or 'G' (green) appeared centred on the screen cueing participants what colour to imagine. The cue appeared for 1s followed by an imagery period of 7 seconds during which only the fixation dot was visible, after which participants rated the vividness of their imagery with a key press (1=lowest, 4=highest). After this, the binocular rivalry display appeared for 0.65s, and participants indicated which colour dominated (1=red, 2=green, 3=mixed). We calculated the strength of individuals' imagery (%primed) as the ratio between the number of primed responses (nprimed) and the resultant number of trials after subtracting mock (nmock) and non-mock mixed response trials (nmixed) from the total number of trials.

fMRI

The stimuli and experiment procedure for our fMRI experiment are as specified in experiment 1 of the 2014 study[1] and as specified in the 2020 study[2] unless reported otherwise. All fMRI stimulation used Presentation (Neurobehavioral Systems). Participants in the scanner wore a blindfold and earbuds for auditory stimulation. Stimuli in the main auditory runs consisted of three natural auditory scenes (a busy road with cars and motorbikes, birds singing and a stream, people talking in a foreign language), which were downloaded from https://www.soundsnap.com, normalised for amplitude (volume), and presented mono, each lasting a total of 12s. The fifth localising run also included a brief "beep" sound. The experiment took place over 6 runs: four auditory runs, a finger tapping run to identify the auditory and motor cortex, and a retinotopic mapping run to identify the regions of interest in the early visual cortex. Each auditory run consisted of all scenes presented 6 times in a pseudo-randomised order with no immediate repetition of the same scene. The scenes lasted 12 seconds followed by 12 seconds of silence (Figure 1A). In the finger tapping run, the paradigm was identical except that we replaced half of the blocks with motor-area-localising blocks during which we instructed participants to press the buttons of a response box in ascending order (right hand only). Beeps signalled participants to start and stop the finger tapping (Figure 1B). The retinotopic mapping run used standard retinotopic polar and eccentricity mappings.[42,43] Before this final run, we told participants to carefully remove their blindfold (something they trained for when we installed them in the scanner).

Data acquisition

We acquired our fMRI data with a 3T Siemens Tim Trio MRI scanner at the Center for Cognitive Neuroimaging (CCNi) of the University of Glasgow. We replicated the acquisition settings (TR = 2s, 222 volumes, TE = 30 ms, resolution: 2.5x2.5x2.5 mm, 35 slices, flip angle 77°, iPAT factor 2) for the main protocol as used in the 2014 study.[1] Additionally, we acquired standard retinotopic and eccentricity data (TR=1000ms, 795 volumes TE=55ms, resolution: 3x3x3 mm, 18 slices) and anatomical data (MPRAGE).

The fMRI data from control participants, obtained by Vetter et al.,[1] was also acquired using the 3T Siemens Tim Trio MRI scanner at the CCNi of the University of Glasgow and, as mentioned, deployed the acquisition settings (apart from an upgrade in head coil to 32-channel from 12-channel). The data from the blind participants, obtained by Vetter et al., was acquired using a 3T General Electric MRI scanner at the Hebrew University of Jerusalem. Acquisition settings differed slightly (TR = 2 s, TE = 30 ms, resolution: 3.2 × 3.2 × 2.5 mm voxels, 35 slices, flip angle: 77°, iPAT factor = 2) from those used in our other two studies. Control analyses were performed showing that blood-oxygen-level-dependent (BOLD) responses were comparable between sites (see Figure.S4 in Vetter et al.[1,2]).''

QUANTIFICATION AND STATISTICAL ANALYSIS

Imagery questionnaires, behavioural testing, and imagery priming results

Participants scored low in both the VVIQ[8,9] (Mean=20.40, SD=4.29) and the SUIS[13] (Mean=19.13, SD=4.07), indicating that they experience low voluntary and spontaneous forms of imagery. In addition to the imagery questionnaires, we performed a behavioural test prior to brain imaging, that capitalises on the fact that individuals with aphantasia cannot be primed through imagery in a binocular rivalry paradigm.[14,44,45] The behavioural test revealed a statistically significant difference in median imagery priming scores between our participants (Median=51.30, SD=7.31) and control data obtained from Bergmann et al.'s typical imagers (Median=56.45, SD=9.79; $\chi^2(1) = 8.76$, $p < 0.01$).[15] An additional Wilcoxon test reported that imagery scores in the controls were significantly different from 50% ($V = 3307$, $p < 0.001$), indicating decisional priming. By contrast, a Wilcoxon test reported that imagery scores in aphantasic participants were not significantly different from 50% ($V=145$, $p>0.05$).

Please cite this article in press as: Montabes de la Cruz et al., Decoding sound content in the early visual cortex of aphantasic participants, Current Biology (2024), https://doi.org/10.1016/j.cub.2024.09.008

Pre-processing, cortical surface reconstruction, and region-of-interest definition

We used BrainVoyager (version 22.4)[33] to perform pre-processing and univariate analysis. Anatomical pre-processing involved the conversion of files from the scanner format into DICOM (performed automatically by BrainVoyager), inhomogeneity correction, ACPC alignment, alignment into Talaraich space, segmentation, and the creation of a mesh for the brain. Cortical reconstruction used BrainVoyager's automatic segmentation for all participants except 3, who required manual segmentations. Functional pre-processing included slice scan time correction, temporal high-pass filter and 3D rigid motion correction (no spatial smoothing). We identified early visual cortices and their eccentricity subdivisions from the retinotopic polar and eccentricity mapping trials.[42,43] We identified the auditory cortex as the area in the lateral sulcus with peak activation for the GLM contrast Sound Stimulation > Baseline from the finger tapping run, and the left motor cortex as the somatosensory area exhibiting peak activation for the GLM contrast Button Pressing > Baseline. In order to map the three cortices, we created a surface map of activity that allowed us to identify them and define the ROI. We combined visual and auditory ROIs across hemispheres. The mean number of vertices across all subjects were as follows: V1: 4950 (SEM=592), V2: 5353 (SEM=764), V3: 4832 (SEM=668), all early visual cortex: 15117 (SEM=1985), auditory cortex ROI: 567 (SEM=65), motor cortex ROI: 892 (SEM=151).

Univariate analysis

We ran a whole-brain random effects GLM analysis for the contrast sound>baseline. The resulting maps were projected on a cortical surface (inflated brain of one participant). Additionally, we extracted mean beta values for each area and sound scene (people talking scene, forest scene and traffic scene). We fitted separate linear mixed models (R + lme4 package) per area to predict beta scores per sound category. The models included sound category as their main effect, and an individual intercept as a random effect.

Multivariate pattern classification analysis

We estimated single block beta weights for all ROI vertices during sound stimulation time courses from the main auditory runs. We then z-scored these beta-weights to fit a linear support vector machine classification algorithm[34] using a one-versus-one classification scheme[46] for each of the three combinations of sounds (people vs forest, people vs traffic and traffic vs forest). We used a leave-one-run-out cross-validation procedure, yielding four classification folds which we averaged to obtain our final classification accuracy. To determine statistical significance, we compared accuracies across 1000 random permutations.

Bayes factors of classification accuracies

We calculated the Bayes factors (BFs) for our decoding accuracies by first establishing probability densities given chance performance level *P(D*|H0*)* with *H0=1/3* and a true effect *P(D|H1)*. For the early visual areas and motor cortex, we defined *H1*=0.4 (a reasonable value given our results in the EVC from the 3 datasets), and for the auditory cortex, we defined *H1*=0.8 to account for the higher expected decoding accuracy for auditory scenes in the auditory cortex vs EVC.

In both cases, we modelled the probability densities using a normal distribution with the mean set to either H0 or H1 and the standard deviation of the area's permutation distribution (see Table S1). The BF is then the ratio of the likelihood of observing the area's classification accuracy D given H0 or H1 (with H0 in numerator so that BF > 1 in case of support for H0):

$$BF = \frac{P(D|H_0)}{P(D|H_1)}$$

Eccentricity modelling in EVC across populations

We fitted a linear mixed model (R + lme4 package) to predict classification accuracy scores. The model included group (aphantasia, control and congenitally blind group), retinotopic area (V1, V2, V3), eccentricity (as a numerical factor, where 0=fovea, 1=periphery and 2=far periphery) and the interaction between group, retinotopic area and eccentricity as its fixed effects. The model also included individual intercept and eccentricity slope as random effects.

Whole brain searchlight analysis

For each subject, we computed a voxelwise searchlight analysis in volume space using a linear support vector machine classifier to decode sound pairs. We defined a searchlight for each voxel coordinate using a radius of 2 units. Voxels selected by each searchlight were determined to be part of the brain by first being part of the set of voxels derived from a full brain region of interest, and secondly by having an average time course BOLD value, or global strength value, greater than 100. We computed decoding performance for each searchlight and stored a cross-validated (leave one run out) average value for each voxel coordinate. This resulted in three maps for each subject.