Electrical stimulation of the superior temporal gyrus evokes rapid responses in human visual cortex

Emily C. Cunningham[1] & David Brang[1]

[1]Department of Psychology
University of Michigan
530 Church St.
Ann Arbor, MI, 48109

Correspondence regarding this article should be addressed to Emily Cunningham at ecngham@umich.edu

COI statement: The authors have no conflicts of interest to disclose.

Abstract

Sounds produce rapid responses in human visual cortex. Animal research suggests that these responses are enabled by corticocortical projections from auditory to visual cortex. To test for the presence of such pathways in humans, we examined intracranial responses to direct electrical stimulation. Stimulation over auditory cortex elicited rapid responses in early visual cortex (particularly near the visual periphery), providing causal support for corticocortical transmission of cross-modal information.

Early visual cortex is increasingly recognized as a site of rapid audiovisual interaction. For example, sounds can modulate activity in the anterior portion of human V1 within 30ms of onset.[1] Although rapid visual responses to auditory stimuli have been repeatedly demonstrated,[1–4] we know relatively little regarding the pathways by which these signals are communicated in humans. Nonhuman primate tracing studies point to several possibilities, including subcortical, corticothalamocortical, and direct corticocortical transmission[5–10]. Of these candidates, the most interest/debate surrounds the possibility that human visual cortex may receive direct anatomical inputs from auditory cortex.

The features of human sound-evoked visual responses are broadly consistent with the existence of such pathways. Their distribution across visual cortex mirrors nonhuman primate anatomy, and their speed of onset suggests rapid communication. Human structural connectivity also provides some support for corticocortical communication.[11,12] However, these observations are indirect, and given the sparse nature of the projections identified in macaques, the question of how these pathways extend to humans remains unresolved (similar responses could be produced, for example, by tectopulvinar transmission[10]).

Here, we provide a causal evaluation of the hypothesis that rapid sound-evoked visual responses originate in auditory cortex. Specifically, we examine how human visual cortex responds to single-pulse electrical stimulation of regions in and around auditory cortex using intracranial EEG (iEEG) recordings from patients with epilepsy (open data from Van Blooijs et al., 2023[13]). This technique provides an estimate of effective connectivity between cortical areas through the distribution/latency of evoked responses (corticocortical evoked potentials, or CCEPs[14]) following current passed between adjacent intracranial contacts. CCEPs typically consist of a rapid early deflection (<50ms) followed by a subsequent slower wave (~80-250ms; Figure 1A). The first-emerging peak in the signal ('N1') is thought to result from corticocortical signal propagation,[14–16] and its distribution/timing may provide insight into the presence of directional, corticocortical projections from auditory to visual cortex.

Occipital CCEPs following stimulation of the superior temporal gyrus (STG) were examined in 23 patients, with a focus on the presence and latency of N1 deflections occurring within 80ms of stimulation (Figure 1A). Stimulation pairs were grouped according to HCP-MMP atlas location of the center of the pair along the STG (Figure 1A). For summary analyses, recording sites were grouped by location within occipital cortex (V1/V2, V3/V4, hMT+, and 'remaining occipital contacts' or 'other'). Two of these regions, V1/V2 and hMT+ (anatomically defined), reflect hypothesized loci of early cross-modal input[5,6,17,18] and are the subject of detailed examination; the remaining two provide points of comparison. For analyses of response prevalence, null distributions were created from 1000 surrogate datasets based on time-series drawn from pre-stimulation windows and processed identically to stimulation-evoked responses.

Overall, 112/231 occipital sites (across 19/23 patients) exhibited early responses to STG stimulation (Figure 1B-C; $p < .001$ vs. response rate in permuted data). The proportion of responses detected was greatest following stimulation of the posterior STG, and declined to near chance with increasingly anterior stimulation (Figure 1B). Within V1/V2, response rates were greatest following mid-STG stimulation (i.e., when stimulation was centered over likely auditory association areas accessible to contacts placed on the outer surface of mid-STG), declining with more anterior/posterior stimulation (Figure 1D; note that only the reduction in response rate between A4/PBelt and Posterior STG was robust to FDR correction), an effect which appears largely driven by responses in anterior portions of the calcarine sulcus and lingual gyrus (Figure 1C). Within lateral occipital regions near hMT+, the proportion of detected responses per patient matched the overall trend, increasing with proximity as stimulation moved from anterior (6%) to posterior STG (26%; Figure 1D). Over the remaining recording sites, responses were often detected at levels above chance, but did not show clear differentiation by stimulation location (Figure 1D).

In nonhuman primates, auditory projections preferentially target anterior regions of V1 corresponding to representation of the peripheral visual field[5,6]. In this dataset, electrode coverage of multiple locations along V1 provided a rare opportunity to explore the presence of a similar bias in the human CCEP response (Figure 2). Responses were observed in the most anterior portion of the calcarine

sulcus (with an average response rate of 33% across 3 patients with anterior calcarine sites; Figure 2A), and to some extent the occipital pole (9%; below the corresponding chance threshold, Figure 2A). There were no detectable early responses in mid-calcarine locations typically associated with parafoveal processing (Figure 2A). The latency of early anterior visual responses ranged from 18-61ms, with 5/7 peaks emerging within 30ms of stimulation (β = 33 ms, 95% CI [22, 45]; Figure 2B). The latency of occipital polar responses ranged from 34-76ms (β = 45 ms, 95% CI [29, 62]). The spectral representation of CCEPs across V1/V2 sites (Figure 2D) was characterized by transient low-frequency activity together with no detectable change in high-gamma power (Figure 2C). This pattern is consistent with modulatory effects produced by feedback-type signal propagation, with no evidence of stimulation-elicited local neuronal spiking typically indexed by increased high-frequency activity;[19,20] It maps well to the anatomy of primate corticocortical projections, which also have a feedback/modulatory character[21].

Patterns of sound-evoked activity[2] and human/nonhuman primate connectivity[11,17,18] also highlight V5/hMT+ as a likely site of early cross-modal input, particularly from auditory motion-sensitive areas along the planum temporale. This is consistent with our general observation of posterior-dominant STG responses in the broad hMT+ ROI (Figure 1D). To clarify the specificity of responses within this region (which was covered in n = 22 patients), we examined whether the relation between posterior STG stimulation and hMT+ activation persisted in the most highly probable hMT+ sites (Figure 3; Figure S3 displays surrounding sites). Notably, the response to posterior STG stimulation was maintained in these sites, but responses to stimulation from other locations along the STG were comparatively limited (i.e., responses in probable hMT+ showed some selectivity to posterior STG; Figure 3A). The latency of probable hMT+ responses ranged from 18-74ms (β = 44 ms, 95% CI [33 54]; Figure 3B). The spectral content of CCEPs in the greater hMT+ area was, as with V1/V2, largely dominated by low-frequency activity, with little evidence of activity in the high-gamma range (Figure 3C).

Overall, these results provide causal support for the presence of direct pathways from auditory to visual cortex in humans. Early visual cortex responded rapidly (18ms+) to electrical stimulation over mid- and posterior STG, with a distribution that largely matched patterns of connectivity observed in humans (using non-invasive methods) and nonhuman primates. Although functional localizers were not available for these patients, the distribution of stimulation-evoked responses also corresponded well to reports of sound-evoked responses in anterior V1[1,3], and reports of auditory motion-related activation of V5/hMT+ and adjacent cortex[22]. STG stimulation also evoked possible responses over the occipital pole. These responses, which were weaker than those over anterior calcarine (see Figure 2A, subject A), have limited support in nonhuman primate data[5], but are consistent with human DTI[12] and cross-modal effects reported at the occipital pole[23,24], and may warrant additional investigation. This profile would be additionally strengthened with investigation of the effects of stimulation over contacts adjacent to A1 and auditory belt/parabelt regions where many of the visual projections identified in macaques originate[5,6,21].

These results indicate that the rapid response to sound in human visual cortex can occur through corticocortical as opposed to solely subcortical transmission. Functionally, the observation that STG stimulation produced low- but not high-frequency changes in visual activity is also consistent with the suggestion that sounds modulate visual cortical excitability without eliciting action potentials[4,23,25]. However, these data do not rule out thalamocortical and corticothalamocortical contributions to the sound-evoked visual response. Respective contributions of these pathways remain an open question. Although the current dataset contained only ipsilateral stimulation/recording combinations, we expect that early cross-modal corticocortical pathways are restricted to the same hemisphere, given animal work and prior research in humans suggesting primacy for spatially-aligned auditory-visual stimuli[2,26,27].

Early cross-modal interactions mediated by corticocortical projections not only support visual processing in the sighted, but also provide avenues for rapid cortical reorganization following vision loss. The use of CCEPs to examine effective connectivity may provide insight into patterns of plasticity following disease/injury by illustrating where cross-modal corticocortical inputs in the sighted are most prevalent. This is likely most relevant for the late blind, but our results also bear similarities to sound-

evoked responses in the occipital cortex of early or congenitally blind individuals. For example, patterns of auditory scene decoding in the early visual cortex of congenitally blind participants also show a pronounced eccentricity bias (present to a lesser extent in sighted subjects)[28], and early blindness is associated with enhanced recruitment of visual motion-sensitive cortex in response to auditory motion[29,30]. Although the current data was obtained from patients with epilepsy, the compatibility of our observations with reports of sound-evoked responses in other patients[1–4], connectivity in nonpatient samples[11,12], and tractography in nonhuman primates[6,17,18,21] provides a measure of confidence in the results.

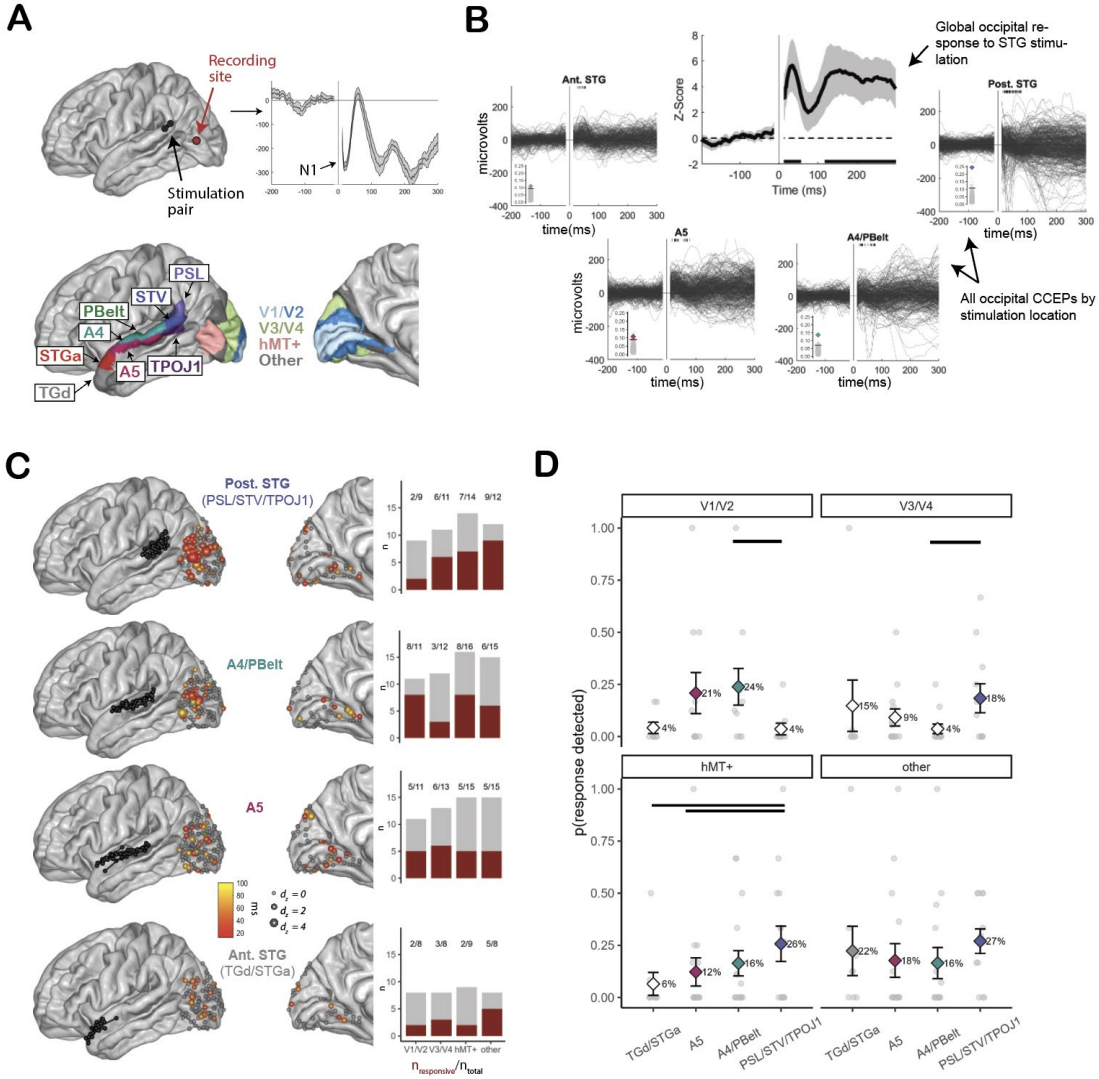

***Figure 1**. Effects of stimulation over STG on responses in occipital cortex. (A) Upper: Illustration of a CCEP response with prominent early negative deflection (N1). Lower: superior temporal and occipital regions of interest. Stimulation locations with limited representation were combined with neighboring locations to create four summary regions: posterior STG, PBelt/A4, A5, and anterior STG (B) Global CCEP response to STG stimulation (center). Black line represents averaged CCEP across all sites & patients (averaged within then across patients); All waveforms with detected responses were rectified*

prior to averaging. Grey ribbon indicates patient-level 95% CI. Black bars at the bottom highlight timepoints at which the observed mean exceeded the 95th percentile of a matched permutation null distribution (FDR-corrected, Benjamini & Hochberg). (outer) all occipital CCEP waveforms split by stimulation location along the STG. Black lines above each plot indicate timepoints at which early peaks were identified. Inset plots: diamonds indicate the average proportion of early responses detected per patient, and grey point-clouds indicate the corresponding proportion of false-positive peaks identified in each of 1000 iterations of shuffled surrogate data; Black lines indicate the 99th percentile of this null distribution. Proportions and surrogate p-values are: Anterior STG: p(resp) = .11, p < .001; A5: p(resp) = .11, p = .002; A4/PBelt: p(resp) = .14, p < .001; Posterior STG: p(resp) = .25, p < .001 (C) Distribution of responsive electrodes as a function of stimulation location along the STG (HCP-MMP atlas). RH electrodes were translated to LH coordinates for visualization. Stimulation pairs within each region are plotted in black. Color indicates latency of the earliest detected peak for a given electrode. Electrodes with no detectable response in the first 80ms are grey. Electrode size is scaled by the effect size of the associated peak (Cohen's d_z). Bar plots indicate the number of patients with at least one responsive electrode (red) relative to the total number of patients (grey) for a given combination of stimulation location and recording region. (D) Average proportion of responses detected per patient (diamonds indicate average; grey dots indicate single patient data) by stimulation and recording region (relative to the total number of CCEP waveforms for a given patient and grouping; see also Figure S4). Error bars are SEM. White diamonds indicate an observed proportion less than the 99th percentile of the associated surrogate null distribution. Black horizontal bars indicate effects of stimulation location on response proportion (after FDR correction across all 24 pairwise combinations). Effects of stimulation location were evaluated using linear mixed-effects models with patient as a random effect; surrogate p-values were computed as the proportion of fixed effects from models fit to null data ≥ the observed fixed effect (V1/V2: PostSTG-A4/PBelt, p = .004, all others n.s.; hMT+: PostSTG-A5, p = .004, PostSTG-AntSTG, p = .002, all others n.s.; V3/V4: PostSTG-A4/PBelt, p = .002, all others n.s.; Other: all n.s.).

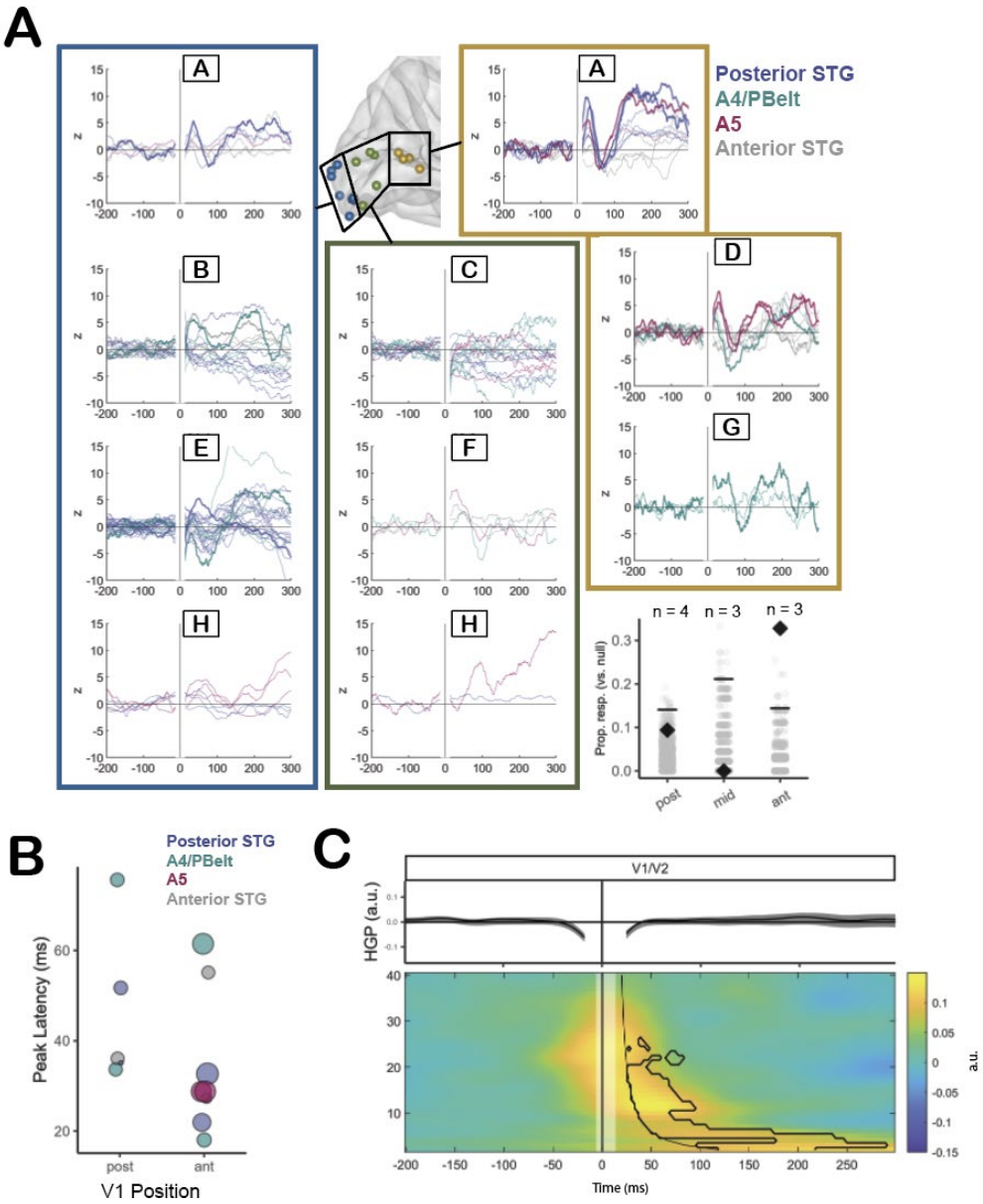

Figure 2. *Examination of V1 responses by eccentricity. (A) All CCEP waveforms from V1 sites. Each plot contains data for a separate patient (identified by letter) within a given region (left = posterior/polar to right = anterior/peripheral). In some cases, as in the top row, a single patient had coverage at multiple points along V1. Bold lines highlight waveforms for which early responses were detected (by criteria: z > 5; peak prominence of ≥ 2). Color of the waveform indicates stimulation location along the STG. Inset: Average proportion of responses detected per patient within each V1 subregion (black diamonds) plotted against the associated surrogate null/chance distribution (grey point clouds). Black bars indicate 99th percentile of the null distribution. (C) Peak latencies for all responsive sites in posterior/anterior V1 (no responses were detected in mid-V1), colored according to stimulation location. Relative size of each point is scaled by effect size at the peak (Cohen's dz; range: 0.39-1.72). Linear mixed-effects models with random effects of subject were applied to derive estimates of peak latency (see main text). (D) Spectral representations of all CCEP waveforms from V1/V2. Upper: high gamma (70-150 Hz) activity time-series. Y-axis ('a.u.') indicates fixed intercept estimates from mixed-effects models of single-trial power for each timepoint, accounting for patient- and electrode-level variance. Error ribbons indicate 95% CI surrounding*

fixed effects (Wald). The area +/- half of the standard deviation of the gaussian window of the longest wavelet is masked (note: over this window, interpolation of the stimulation artifact artificially reduces high-gamma power). Lower: Heat maps depict time-frequency decomposition (2-40 Hz) via Morlet wavelets. Color ('a.u.') indicates fixed intercept estimates from mixed-effects models of single-trial power for each time/frequency point, accounting for patient- and electrode-level variance. Faded central bars indicate the window over which the stimulation artifact was interpolated prior to time-frequency decomposition (-5 to 12ms). For illustrative purposes, black lines indicate FDR-corrected regions showing significant difference from baseline. The curve of the line following stimulation onset highlights the standard deviation of the gaussian window for the wavelet at each frequency (i.e., the region of potential interference from the stimulation artifact, over which statistics were not evaluated).

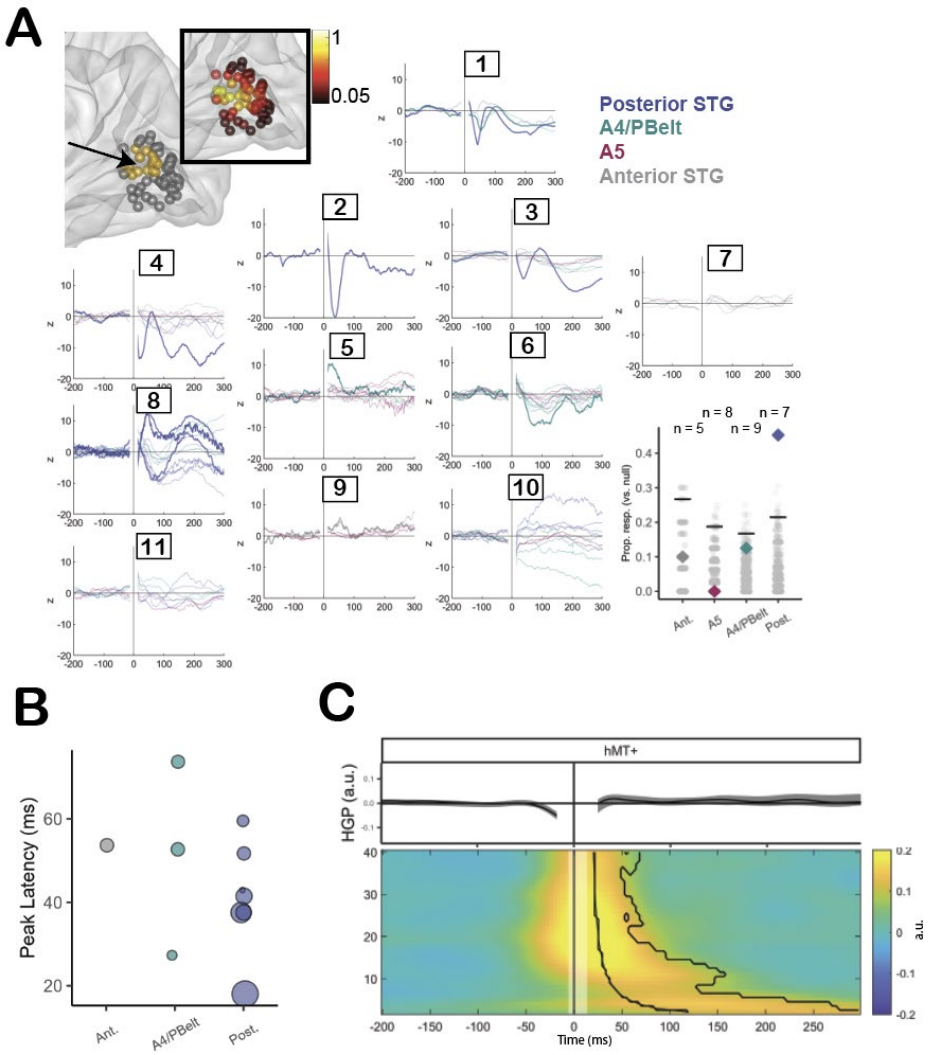

Figure 3. *Examination of responses in the most probable MT/MST sites. (A) All CCEP waveforms from probable hMT+ sites. Electrode locations within the greater ROI are highlighted in the upper left, with inset plot indicating probability of membership. Each subplot contains data for a separate patient (identified by number). Bold lines highlight waveforms for which early responses were detected. Color of the waveform indicates stimulation location along the STG. Lower right: Average proportion of responses detected per patient for each stimulation location (diamonds) plotted against the associated surrogate null/chance distribution (grey point clouds). Number of patients contributing data for each stimulation location is indicated above the plot. Black bars indicate 99th percentile of the null distribution. (B) Peak*

latencies for all responsive sites in probable hMT+ colored according to stimulation location. Relative size of each point is scaled by effect size at the peak (Cohen's dz; range: 0.23-6.07). Linear mixed-effects models with random effects of subject were applied to derive estimates of peak latency (main text). (C) Spectral representations of all CCEP waveforms from the broad hMT+ roi. Upper: high gamma (70-150 Hz) activity time-series. Y-axis ('a.u.') indicates fixed intercept estimates from mixed-effects models of single-trial power for each timepoint, accounting for patient- and electrode-level variance. Error ribbons indicate 95% CI surrounding fixed effects (Wald). The area +/- half of the standard deviation of the gaussian window of the longest wavelet is masked (note: over this window, interpolation of the stimulation artifact artificially reduces high-gamma power). Lower: Heat maps depict time-frequency decomposition (2-40 Hz) via Morlet wavelets. Color ('a.u.') indicates fixed intercept estimates from mixed-effects models of single-trial power for each time/frequency point, accounting for patient- and electrode-level variance. Faded central bars indicate the window over which the stimulation artifact was interpolated prior to time-frequency decomposition (-5 to 12ms). For illustrative purposes, black lines indicate FDR-corrected regions showing significant difference from baseline. The curve of the line following stimulation onset highlights the standard deviation of the gaussian window for the wavelet at each frequency (i.e., the region of potential interference from the stimulation artifact, over which statistics were not evaluated).

References

1. Brang D, Towle VL, Suzuki S, Hillyard SA, Di Tusa S, Dai Z, Tao J, Wu S, Grabowecky M. Peripheral sounds rapidly activate visual cortex: evidence from electrocorticography. Journal of Neurophysiology. 2015 Nov 1;114(5):3023–3028.

2. Plass J, Ahn E, Towle VL, Stacey WC, Wasade VS, Tao J, Wu S, Issa NP, Brang D. Joint encoding of auditory timing and location in visual cortex. J Cogn Neurosci. 2019 Jul;31(7):1002–1017. PMCID: PMC6551606

3. Ferraro S, Van Ackeren MJ, Mai R, Tassi L, Cardinale F, Nigri A, Bruzzone MG, D'Incerti L, Hartmann T, Weisz N, Collignon O. Stereotactic electroencephalography in humans reveals multisensory signal in early visual and auditory cortices. Cortex. 2020 May 1;126:253–264.

4. Mercier MR, Foxe JJ, Fiebelkorn IC, Butler JS, Schwartz TH, Molholm S. Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory integration. NeuroImage. 2013 Oct 1;79:19–29.

5. Falchier A, Clavagnier S, Barone P, Kennedy H. Anatomical Evidence of Multimodal Integration in Primate Striate Cortex. J Neurosci. Society for Neuroscience; 2002 Jul 1;22(13):5749–5759. PMID: 12097528

6. Rockland KS, Ojima H. Multisensory convergence in calcarine visual areas in macaque monkey. International Journal of Psychophysiology. 2003 Oct 1;50(1):19–26.

7. Borra E, Rockland KS. Projections to Early Visual Areas V1 and V2 in the Calcarine Fissure from Parietal Association Areas in the Macaque. Front Neuroanat. Frontiers; 2011 Jun 22

8. Clavagnier S, Falchier A, Kennedy H. Long-distance feedback projections to area V1: Implications for multisensory integration, spatial awareness, and visual consciousness. Cognitive, Affective, & Behavioral Neuroscience. 2004 Jun 1;4(2):117–126.

9. Cappe C, Rouiller EM, Barone P. Multisensory anatomical pathways. Hearing Research. 2009 Dec 1;258(1):28–36.

10. Lyon DC, Nassi JJ, Callaway EM. A Disynaptic Relay from Superior Colliculus to Dorsal Stream Visual Cortex in Macaque Monkey. Neuron. 2010 Jan 28;65(2):270–279.

11. Gurtubay-Antolin A, Battal C, Maffei C, Rezk M, Mattioni S, Jovicich J, Collignon O. Direct Structural Connections between Auditory and Visual Motion-Selective Regions in Humans. J Neurosci. Society for Neuroscience; 2021 Mar 17;41(11):2393–2405. PMID: 33514674

12. Beer AL, Plank T, Greenlee MW. Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. Exp Brain Res. 2011 Sep 1;213(2):299–308.

13. van Blooijs D, van den Boom MA, van der Aar JF, Huiskamp GM, Castegnaro G, Demuru M, Zweiphenning WJEM, van Eijsden P, Miller KJ, Leijten FSS, Hermes D. Developmental trajectory of transmission speed in the human brain. Nat Neurosci. Nature Publishing Group; 2023 Apr;26(4):537–541.

14. Matsumoto R, Nair DR, LaPresto E, Najm I, Bingaman W, Shibasaki H, Lüders HO. Functional connectivity in the human language system: a cortico-cortical evoked potential study. Brain. 2004 Oct;127(10):2316–2330.

15. Conner CR, Ellmore TM, DiSano MA, Pieters TA, Potter AW, Tandon N. Anatomic and electro-physiologic connectivity of the language system: A combined DTI-CCEP study. Computers in Biology and Medicine. 2011 Dec 1;41(12):1100–1109.

16. Matsumoto R, Kunieda T, Nair D. Single pulse electrical stimulation to probe functional and pathological connectivity in epilepsy. Seizure. 2017 Jan 1;44:27–36.

17. Palmer SM, Rosa MGP. Quantitative Analysis of the Corticocortical Projections to the Middle Temporal Area in the Marmoset Monkey: Evolutionary and Functional Implications. Cerebral Cortex. 2006 Sep 1;16(9):1361–1375.

18. Majka P, Rosa MGP, Bai S, Chan JM, Huo BX, Jermakow N, Lin MK, Takahashi YS, Wolkowicz IH, Worthy KH, Rajan R, Reser DH, Wójcik DK, Okano H, Mitra PP. Unidirectional monosynaptic connections from auditory areas to the primary visual cortex in the marmoset monkey. Brain Struct Funct. 2019 Jan 1;224(1):111–131.

19. Bastos AM, Vezoli J, Bosman CA, Schoffelen JM, Oostenveld R, Dowdall JR, De Weerd P, Kennedy H, Fries P. Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. Neuron. Elsevier; 2015 Jan 21;85(2):390–401. PMID: 25556836

20. Buffalo EA, Fries P, Landman R, Buschman TJ, Desimone R. Laminar differences in gamma and alpha coherence in the ventral stream. Proceedings of the National Academy of Sciences. Proceedings of the National Academy of Sciences; 2011 Jul 5;108(27):11262–11267.

21. Falchier A, Cappe C, Barone P, Schroeder CE. Sensory Convergence in Low-Level Cortices. In: Stein BE, editor. The New Handbook of Multisensory Processing. The MIT Press; 2012

22. Rezk M, Cattoir S, Battal C, Occelli V, Mattioni S, Collignon O. Shared Representation of Visual and Auditory Motion Directions in the Human Middle-Temporal Cortex. Current Biology. Elsevier; 2020 Jun 22;30(12):2289-2299.e8. PMID: 32442465

23. Romei V, Murray MM, Cappe C, Thut G. Preperceptual and Stimulus-Selective Enhancement of Low-Level Human Visual Cortex Excitability by Sounds. Current Biology. 2009 Nov 17;19(21):1799–1805.

24. Romei V, Murray MM, Merabet LB, Thut G. Occipital Transcranial Magnetic Stimulation Has Opposing Effects on Visual and Auditory Stimulus Detection: Implications for Multisensory Interactions. J Neurosci. Society for Neuroscience; 2007 Oct 24;27(43):11465–11472. PMID: 17959789

25. Iurilli G, Ghezzi D, Olcese U, Lassi G, Nazzaro C, Tonini R, Tucci V, Benfenati F, Medini P. Sound-Driven Synaptic Inhibition in Primary Visual Cortex. Neuron. Elsevier; 2012 Feb 23;73(4):814–828. PMID: 22365553

26. Wallace MT, Wilkinson LK, Stein BE. Representation and integration of multiple sensory inputs in primate superior colliculus. Journal of Neurophysiology. American Physiological Society Bethesda, MD; 1996

27. McDonald JJ, Teder-Sälejärvi WA, Hillyard SA. Involuntary orienting to sound improves visual perception. Nature. Nature Publishing Group; 2000 Oct;407(6806):906–908.

28. Vetter P, Bola Ł, Reich L, Bennett M, Muckli L, Amedi A. Decoding Natural Sounds in Early “Visual” Cortex of Congenitally Blind Individuals. Current Biology. Elsevier; 2020 Aug 3;30(15):3039-3044.e2. PMID: 32559449

29. Poirier C, Collignon O, Scheiber C, Renier L, Vanlierde A, Tranduy D, Veraart C, De Volder AG. Auditory motion perception activates visual motion areas in early blind subjects. NeuroImage. 2006 May 15;31(1):279–285.

30. Strnad L, Peelen MV, Bedny M, Caramazza A. Multivoxel Pattern Analysis Reveals Auditory Motion Information in MT+ of Both Congenitally Blind and Sighted Individuals. PLOS ONE. Public Library of Science; 2013 Apr 30;8(4):e63198.

Method

Participants

Data were drawn from the open repository associated with van Blooijs et al., 2023. A subset (N = 23; mean age = 25 years, range 8-51) of the original lifespan sample of 74 participants had combinations of STG stimulation and occipital recording sites that were relevant for the current question (see Exclusion/Inclusion criteria below), but which were not considered in the initial paper. In all cases, sites were positioned ipsilaterally, with the majority on the left hemisphere (17/23; Figure S1).

Exclusion/Inclusion criteria. Data were screened to identify patients with some combination of STG and occipital electrode coverage, defined in terms of Destrieux atlas labels from individual cortical parcellations (Destrieux et al., 2010). Patients with relevant coverage were included if they had (a) at least one stimulation pair in which both electrodes were located on the STG, and (b) at least 10 trials for analysis for at least one relevant stimulation pair.

Data Acquisition

For details of data acquisition, readers are referred to van Blooijs et al. (2023). In brief, single-pulse electrical stimulation was conducted as iEEG data were recorded from subdural electrode strips/grids with 1 cm spacing. Data was initially acquired at a sampling rate of between 512-2048 Hz. For a given pair of adjacent electrodes, stimulation (pulse width: 1ms, current intensity: 8mA) was applied up to 10 times at intervals of 5s.

Processing & Analysis

Selection of stimulation pairs and recording sites. For patients meeting inclusion criteria, stimulation pairs were included in the following analyses if both members of a pair were located over the STG, or if one member was located on the STG and the other was located on the inferior portion of the supramarginal gyrus. The region 'PSL' from the Glasser (2016) parcellation was used to define the upper border for inclusion of supramarginal electrodes. Pairs that crossed the superior temporal sulcus or lateral fissure were not included. Stimulation pairs were also excluded if they contained any channels marked for exclusion by the criteria described in van Blooijs et al. (2023), or if the total number of trials for a pair was less than 10. All trials for a given stimulation pair were included, except in rare instances in which stimulation occurred within 3s of another trial. Occipital recording sites were included if their locations fell within the set of Destrieux atlas labels {G&S_occipital_inf, G_cuneus, G_oc-temp_med-Lingual, G_occipital_middle, G_occipital_sup, Pole_occipital, S_calcarine, S_oc_middle&Lunatus, S_oc_sup&transversal, S_occipital_ant, S_parieto_occipital}. Additional exclusion criteria based on data quality are described below.

Preprocessing and measurement of evoked responses. As in Van Blooijs et al (2023), no modification of the continuous signal was performed (i.e., data were neither rereferenced nor filtered). For measurement of evoked responses, timepoints between -12 and 12ms were excluded from analysis to limit contamination from stimulation artifacts. CCEPs were evaluated relative to a prestimulus average baseline of [-200 ms to -12 ms]. For peak latency measurement, the averaged ERP for a given combination of electrode site and stimulation pair was transformed by z-scoring the waveform relative to the mean and standard deviation over the baseline window (after the approach employed by Trebaul et al., 2018; Lemarechal et al., 2022). Peaks were identified as the first minimum/maximum in the interval [12 to 250ms] with an absolute value exceeding z = 5 and a minimum peak prominence of 2 (i.e. peaks were required to have a drop of z = 2 on either side before encountering a neighboring peak of greater amplitude). Electrodes were considered to show an early response if the CCEP waveform contained a detectable peak in the range 12-80ms following stimulation. Note that the use of an 80ms window risks

inclusion of some peaks later than the N1 component (which typically peaks in the first 50ms). However, extending the range of inquiry to 80ms ensures more complete description of early responses.

Saturated responses (defined by post-stimulation amplitudes exceeding 2000 microvolts) were excluded from analysis (n = 3 electrodes, 1 patient), as were responses that showed extreme time-locked line-noise contamination in the averaged waveform (defined as having a mean power spectral density over the range 50-60 Hz, prior to stimulation, that was greater than 10x the median value over all stimulation pairs and electrode sites; n = 62 responses over 8 patients; this exclusion criterion was data-driven and implemented to reduce the occurrence of false-positive 'peaks' resulting from line noise). Overall, after these data quality exclusions, 1305 combinations of stimulation pair and recording site (231 total recording sites) were considered for analysis. Included stimulation pairs and recording sites for each patient are presented in Figure S1.

Given that the z-scored waveforms used to compute peak latency are scored relative to variation in the *averaged* waveform over the baseline window, the z-scores alone contain no information regarding variance in the waveform from trial to trial. To provide a coarse index of stability/variation in amplitude across repeated stimulations, peak latency values were supplemented with Cohen's d_z values computed for peak amplitude (computed by dividing peak amplitude by the standard deviation of the waveform across trials at that timepoint).

Two permutation approaches were adopted to provide chance baselines against which to compare the observed data. First, for time series evaluation, the following procedure was conducted to provide a simple summary assessment and visualization of the occipital response to STG stimulation. First, 1000 null datasets were generated using a sign-based permutation approach, in which the values for half of the trials associated with each CCEP waveform were randomly inverted prior to averaging. For each null dataset, CCEP waveforms were generated and peaks evaluated exactly as in the observed data. Then, for both the null and observed data, to account for the fact that CCEP polarity can vary by stimulation pair and recording site, all waveforms for which a negative early peak was detected were inverted to place all peaks (for both null and comparison distributions) in a consistent direction. All waveforms across all electrodes (regardless of whether a response was detected) were then averaged within-subject and across subjects to generate a summary CCEP reflecting the entire occipital ROI. Surrogate p-values were derived from comparisons of the observed deviation from baseline and the associated permutation distribution at a given timepoint (FDR-corrected, Benjamini & Hochberg, 1995).

Second, to generate a chance baseline that best reflects the expected false-positive rate in the absence of stimulation, we generated a comparison distribution out of 1000 repetitions of the following: For each CCEP time-series, we extracted a matched null time-series centered around a point within the 2 seconds prior to that trial (selected randomly for each repetition). These null time-series were then processed in the same way as the observed data described above, resulting in 1000 surrogate datasets identical to the main dataset except for the use of null data. In this way, it was possible to evaluate expected false positive rates assuming no effect of stimulation for every observed statistic. Observed values were evaluated against this null distribution, and surrogate p-values were computed as the proportion of null values matching or exceeding the magnitude of the observed statistic. The 99th percentile of the null distribution was used as a benchmark for evaluating response rates in a given region.

Summary statistics by stimulation location and region. Stimulation pairs were grouped by the HCP-MMP (Glasser et al., 2016) parcellation label associated with the nearest coordinate in Freesurfer average space to the center of the pair. Stimulation locations with limited representation were combined

with neighboring locations to create four summary regions: posterior STG, PBelt/A4, A5, and anterior STG (see Figure S2 for a full breakdown of the distribution of responses for all STG labels).

For summary analyses over visual locations, we defined sites near probable V5/hMT+ as those where nearest coordinates in Freesurfer fsaverage space were positioned within the bounds of a mask drawn by applying a >5% threshold to the probabilistic atlas developed by Huang et al., 2019 (version 'thr3_snrH'). Remaining locations in and around V1/V2 and V3/V4 were defined by proximity of nearest coordinates in Freesurfer average space to HCP-MMP (Glasser et al., 2016) parcellation labels (Figure 1).

For comparisons of response rates (N1 prevalence) by stimulation location within each visual location, linear mixed-effects models (maximum likelihood estimation) were generated for each pair of stimulation locations within a region using the lme4 package in R (Bates et al., 2015), with random intercepts for patient ID to address situations in which the same patient received stimulation from multiple STG locations. The fixed effect of stimulation location was compared against the fixed effects derived from identical models fit to the null data. The surrogate p-values obtained from these comparisons (proportion of null responses matching or exceeding the magnitude of the observed fixed effect) were corrected across the full set of 24 comparisons (6 comparisons within each of four visual ROIs; FDR, Benjamini & Hochberg, 1995).

Exploratory analyses of V1 and hMT+. To explore responses within early visual cortex as a function of eccentricity, V1 sites were divided into anterior/middle/posterior regions, split based on fsaverage y-coordinate (divided at -94 and -80, chosen by visual inspection). To explore specificity of hMT+ responses, the initial hMT+ ROI was reduced to highlight only those locations with maximum probability MT/MST labels on the HCP-MMP atlas (Glasser et al., 2016).

To summarize peak latency within these locations, accounting for variance in number of responsive sites across patients, linear mixed-effects models with random effects of patient identity were generated for each region. These simple models contained the random effects and a fixed intercept, the value of which was used as an estimate of the population mean peak latency (Bates et al., 2015). Confidence intervals around the fixed effect (95%) computed using the Wald method provide estimates of uncertainty.

Spectral analyses. For the detailed descriptions of V1/V2 and hMT+ responses, spectral decomposition was performed to characterize the frequency content of the signal over all CCEPs in the region. The intent of this decomposition was descriptive. Trials were epoched from -3 to 3s to allow a sufficient window for event-related spectral perturbation (ERSP) computation. The stimulation artifact was removed by interpolating the time-series between -5 and 12 ms for each trial with piecewise cubic hermite interpolation. Single-trial spectrograms were computed by convolving the iEEG time series with complex Morlet wavelets (for lower frequencies: 39 center frequencies linearly spaced between 2-40 Hz, with number of cycles per frequency increasing linearly from 3 to 12; for high-gamma power: 16 center frequencies linearly spaced between 75 and 150 Hz, with number of cycles per frequency increasing linearly from 12 to 24). A time step of ~4ms was applied to produce time-frequency plots sampled at 256 Hz. Following convolution, power was extracted as squared amplitude for each trial, timepoint, and frequency. Spectrograms were baseline-corrected and normalized using the 'vssum' approach: for each timepoint the average power over time and trials in the interval [-500 to -200ms] at that frequency was subtracted, and values were then normalized by dividing by the sum of the value at that timepoint and the associated baseline value. The resulting values are bounded between -1 and 1 and scaled such that +/- 0.33 indicates a 2x increase/decrease in power from baseline. High-gamma power values were averaged across all frequencies to produce a single time-series representation of high-frequency activity.

Single-trial spectral information was modeled with a mixed-effects approach to generate fixed-effect estimates of power relative to baseline (used to create time-frequency visualizations), and associated confidence intervals (used in representations of high-gamma time series) for each time/frequency point. For these models (with trial nested within electrode nested within subject), the random effects structure included crossed random effects of electrode, run (for patients with multiple runs), and stimulation pair. Confidence intervals around the fixed effect (95%) were computed using the Wald method for high-gamma time-series. For time-frequency maps, to visualize uncertainty, p-values extracted using lmerTest were evaluated over the post-stimulation interval (FDR corrected, Benjamini-Hochberg). To minimize contamination by the interpolated stimulation artifact, these analyses were conducted over the interval starting at the end of the interpolation window (12ms) plus half the width of the standard deviation of the gaussian window applied at a given frequency. Note that interpolation of the stimulation artifact produces an artificial reduction in high-gamma power in the peristimulus interval.