Check for updates

Functional selectivity for face processing in the temporal voice area of early deaf individuals

Stefania Benetti[a,1], Markus J. van Ackeren[a], Giuseppe Rabini[a], Joshua Zonca[a], Valentina Foa[a], Francesca Baruffaldi[b], Mohamed Rezk[a], Francesco Pavani[a,c,d], Bruno Rossion[e,f], and Olivier Collignon[a,e,f,1]

[a]Center for Mind/Brain Studies, University of Trento, 38123 Trento, Italy; [b]Sezione di Trento, Ente Nazionale Sordi, 38122 Trento, Italy; [c]Department of Psychology and Cognitive Sciences, University of Trento, 38068 Rovereto, Italy; [d]Integrative, Multisensory, Perception, Action and Cognition Team, Centre de Recherche en Neurosciences de Lyon, 69676 BRON Cedex, Lyon, France; [e]Institute of Research in Psychology, University of Louvain, 1348 Louvain-la-Neuve, Belgium; and [f]Institute of NeuroScience, University of Louvain, 1348 Louvain-la-Neuve, Belgium

Brain systems supporting face and voice processing both contribute to the extraction of important information for social interaction (e.g., person identity). How does the brain reorganize when one of these channels is absent? Here, we explore this question by combining behavioral and multimodal neuroimaging measures (magneto-encephalography and functional imaging) in a group of early deaf humans. We show enhanced selective neural response for faces and for individual face coding in a specific region of the auditory cortex that is typically specialized for voice perception in hearing individuals. In this region, selectivity to face signals emerges early in the visual processing hierarchy, shortly after typical face-selective responses in the ventral visual pathway. Functional and effective connectivity analyses suggest reorganization in long-range connections from early visual areas to the face-selective temporal area in individuals with early and profound deafness. Altogether, these observations demonstrate that regions that typically specialize for voice processing in the hearing brain preferentially reorganize for face processing in born-deaf people. Our results support the idea that cross-modal plasticity in the case of early sensory deprivation relates to the original functional specialization of the reorganized brain regions.

cross-modal plasticity | deafness | modularity | ventral stream | identity processing

T he human brain is endowed with the fundamental ability to adapt its neural circuits in response to experience. Sensory deprivation has long been championed as a model to test how experience interacts with intrinsic constraints to shape functional brain organization. In particular, decades of neuroscientific research have gathered compelling evidence that blindness and deafness are associated with cross-modal recruitment of the sensory-deprived cortices (1). For instance, in early deaf individuals, visual and tactile stimuli induce responses in regions of the cerebral cortex that are sensitive primarily to sounds in the typical hearing brain (2, 3).

Animal models of congenital and early deafness suggest that specific visual functions are relocated to discrete regions of the reorganized cortex and that this functional preference in cross-modal recruitment supports superior visual performance. For instance, superior visual motion detection is selectively altered in deaf cats when a portion of the dorsal auditory cortex, specialized for auditory motion processing in the hearing cat, is transiently deactivated (4). These results suggest that cross-modal plasticity associated with early auditory deprivation follows organizational principles that maintain the functional specialization of the colonized brain regions. In humans, however, there is only limited evidence that specific nonauditory inputs are differentially localized to discrete portions of the auditory-deprived cortices. For example, Bola et al. have recently reported, in deaf individuals, cross-modal activations for visual rhythm discrimination in the posterior-lateral and associative auditory regions that are recruited by auditory rhythm discrimination in hearing individuals (5). However, the observed cross-modal recruitment encompassed an extended portion of these temporal regions, which were found activated also by other

visual and somatosensory stimuli and tasks in previous studies (2, 3). Moreover, it remains unclear whether specific reorganization of the auditory cortex contributes to the superior visual abilities documented in the early deaf humans (6). These issues are of translational relevance because auditory reafferentation in the deaf is now possible through cochlear implants and cross-modal recruitment of the temporal cortex is argued to be partly responsible for the high variability, in speech comprehension and literacy outcomes (7), which still poses major clinical challenges.

To address these issues, we tested whether, in early deaf individuals, face perception selectively recruits discrete regions of the temporal cortex that typically respond to voices in hearing people. Moreover, we explored whether such putative face-selective cross-modal recruitment is related to superior face perception in the early deaf. We used face perception as a model based on its high relevant social and linguistic valence for deaf individuals and the suggestion that auditory deprivation might be associated with superior face-processing abilities (8). Recently, it was demonstrated that both linguistic (9) and nonlinguistic (10) facial information remaps to temporal regions in postlingually deaf individuals. In early deaf individuals, we expected to find face-selective responses in the middle and ventrolateral portion of the auditory cortex, a region showing high sensitivity to vocal acoustic information in hearing individuals: namely the "temporal voice-selective area" (TVA) (11). This hypothesis is notably based on the observation that facial and vocal signals are integrated in lateral belt regions of

Significance

Here, we show that deaf individuals activate a specific and discrete subregion of the temporal cortex, typically selective to voices in hearing people, for visual face processing. This reorganized "voice" region participates in face identity processing and responds selectively to faces early in time, suggesting that this area becomes an integral part of the face network in early deaf individuals. Observing that face processing selectively colonizes a region of the hearing brain that is functionally related to identity processing evidences the intrinsic constraints imposed to cross-modal plasticity. Our work therefore supports the view that, even if brain components modify their sensory tuning in case of deprivation, they maintain a relation to the computational structure of the problems they solve.

the monkey temporal cortex (12). Moreover, there is evidence for functional interactions between this portion of the TVA and the face-selective area of the ventral visual stream in the middle lateral fusiform gyrus [the fusiform face area (FFA)] (13) during person recognition in hearing individuals (14), and of direct structural connections between these regions in hearing individuals (15). To further characterize the potential role of reorganized temporal cortical regions in face perception, we also investigated whether these regions support face identity discrimination by means of a repetition–suppression experiment in functional magnetic resonance imaging (16). Next, we investigated the time course of putative TVA activation during face perception by reconstructing virtual time series from magneto-encephalographic (MEG) recordings while subjects viewed images of faces and houses. We predicted that, if deaf TVA has an active role in face perception, category selectivity should be observed close in time to the first selective response to faces in the fusiform gyrus: i.e., between 100 and 200 ms (17). Finally, we examined the role of long-range corticocortical functional connectivity in mediating the potential cross-modal reorganization of TVA in the deaf.

Results

Experiment 1: Face Perception Selectively Recruits Right TVA in Early Deaf Compared with Hearing Individuals.

To test whether face perception specifically recruits auditory voice-selective temporal regions in the deaf group ($n = 15$), we functionally localized (*i*) the TVA in a group of hearing controls ($n =15$) with an fMRI voice localizer and (*ii*) the face-selective network in each group [i.e., hearing controls = 16; hearing users of the Italian Sign Language (LIS) = 15; and deaf individuals = 15] with an fMRI face localizer contrasting full-front images of faces and houses matched for low-level features like color, contrast, and spatial frequencies (*Materials and Methods*). A group of hearing users of the Italian Sign Language was included in the experiment to control for the potential confounding effect of exposure to visual language. Consistent with previous studies of face (13) and voice (11) perception, face-selective responses were observed primarily in the midlateral fusiform gyri bilaterally, as well as in the right posterior superior temporal sulcus (pSTS) across the three groups (*SI Appendix*, Fig. S1 and Table 1) whereas voice-selective responses were observed in the midlateral portion of the superior temporal gyrus (mid-STG) and the midupper bank of the STS (mid-STS) in the hearing control group (*SI Appendix*, Fig. S2).

When selective neural responses to face perception were compared among the three groups, enhanced face selectivity was observed in the right midlateral STG, extending ventrally to the midupper bank of the STS [Montreal Neurological Institute (MNI) coordinates (62 –18 2)] in the deaf group compared with both the hearing and the hearing-LIS groups (Fig. 1 *A* and *B* and

Table 1. Regional responses for the main effect of face condition in each group and differences between the three groups

Area	Cluster size	X_{mm}	Y_{mm}	Z_{mm}	Z	df	P_{FWE}
Hearing controls faces > houses						15	
R fusiform gyrus (lateral)	374	44	−50	−16	4.89		0.017*
R superior temporal gyrus/sulcus (posterior)	809	52	−42	16	4.74		<0.001
R middle frontal gyrus	1,758	44	8	30	4.33		<0.001
L fusiform gyrus (lateral)	97	−42	−52	−20	4.29		0.498
Hearing-LIS faces > houses						14	
R superior temporal gyrus/sulcus (posterior)	1,665	52	−44	10	5.64		<0.001*
R middle frontal gyrus	3,596	34	4	44	5.48		<0.001*
R inferior frontal gyrus	S.C.	48	14	32	5.45		
R inferior parietal gyrus	1,659	36	−52	48	5.45		<0.001*
R fusiform gyrus (lateral)	46[†]	44	−52	−18	3.52		0.826
L middle frontal gyrus	1,013	−40	4	40	5.13		<0.001*
L fusiform gyrus (lateral)	120	−40	−46	−18	4.02		0.282
R/L superior frontal gyrus	728	2	20	52	4.78		<0.001
R/L superior frontal gyrus	728	2	20	52	4.78		<0.001
Deaf faces > houses						14	
R middle frontal gyrus	1,772	42	2	28	4.84		<0.001*
R inferior temporal gyrus	845	50	−60	−10	4.04		<0.001
R fusiform gyrus (lateral)	S.C.	48	−56	−18	3.54		
R superior temporal gyrus/sulcus (posterior)	S.C.	50	−40	14	4.01		
R superior temporal gyrus/sulcus (middle)	64	54	−24	−4	3.80		0.005[‡]
R thalamus (posterior)	245	10	−24	10	4.02		0.042
L fusiform gyrus (lateral)	209	−40	−66	−18	3.90		0.071
R putamen	329	28	0	4	4.49		0.013
L middle frontal gyrus	420	−44	26	30	3.95		0.004
Deaf > hearing controls ∩ hearing-LIS faces > houses						3,44	
R superior temporal gyrus/sulcus (middle)	73	62	−18	2	3.86		0.006[‡]
Deaf > hearing controls faces > houses						30	
R superior temporal gyrus/sulcus (middle)	167	62	−18	4	3.77		0.001[‡]
L superior temporal gyrus/sulcus (middle)	60	−64	−24	10	3.64		0.007[‡]
Deaf > hearing-LIS faces > houses						29	
R superior temporal gyrus/sulcus (middle)	73	62	−18	2	3.86		0.006[‡]

Significance corrections are reported at the cluster level; cluster size threshold = 50. df, degrees of freedom; FWE, family-wise error; L, left; R, right; S.C., same cluster.
*Brain activations significant after FWE voxel correction over the whole brain.
[†]Cluster size <50.
[‡]Brain activation significant after FWE voxel correction over a small spherical volume (25-mm radius) at peak coordinates for right and left hearing TVA.

Table 1). The location of this selective response strikingly overlapped with the superior portion of the right TVA as functionally defined in our hearing control group (Fig. 1 C and D). Face selectivity was additionally observed in the left dorsal STG posterior to TVA [MNI coordinates (−64 −28 8)] when the deaf and hearing control group where compared; however, no differences were detected in this region when the deaf and hearing control groups were, respectively, compared with hearing-LIS users. To further describe the preferential face response observed in the right temporal cortex, we extracted individual measures of estimated activity (beta weights) in response to faces and houses from the right TVA as independently localized in the hearing groups. In these regions, an analysis of variance revealed an interaction effect [$F_{(category \times group)} = 16.18$, $P < 0.001$, $\eta^2 = 0.269$], confirming increased face-selective response in the right mid-STG/STS of deaf individuals compared with both the hearing controls and hearing LIS users [$t_{(deaf > hearing)} = 3.996$, $P < 0.001$; Cohen's $d = 1.436$; $t_{(deaf > hearing-LIS)} = 3.907$, $P < 0.001$, Cohen's $d = 7.549$] (Fig. 1D). Although no face selectivity was revealed—at the whole brain level and with small volume correction (SVC)—in the left temporal cortex of deaf individuals, we further explored the individual responses in left mid-TVA for completeness. Cross-modal face selectivity was also revealed in this region in the deaf although the interindividual variability within this group was larger and the face-selective response was weaker (*SI Appendix, Supporting Information* and Fig. S4). In contrast to the preferential response observed for faces, no temporal region showed group differences

for house-selective responses (Table 1 and *SI Appendix*, Fig. S1). Hereafter, we focus on the right temporal region showing robust face-selective recruitment in the deaf and refer to it as the deaf temporal face area (dTFA).

At the behavioral level, performance in a well-known and validated neuropsychological test of individual face matching (the Benton Facial Recognition Test) (18) and a delayed recognition of facial identities seen in the scanner were combined in a composite face-recognition measure in each group. This composite score was computed to achieve a more stable and comprehensive measure of the underlying face processing abilities (19).When the three groups were compared on face-processing ability, the deaf group significantly outperformed the hearing group (t = 3.066, P = 0.012, Cohen's d = 1.048) (Fig. 1E) but not the hearing-LIS group, which also performed better than the hearing group (t = 3.080, P = 0.011, Cohen's d = 1.179) (Fig. 1E). This result is consistent with previous observations suggesting that both auditory deprivation and use of sign language lead to a superior ability to process face information (20). To determine whether there was a relationship between face-selective recruitment of the dTFA and face perception, we compared interindividual differences in face-selective responses with corresponding variations on the composite measure of face recognition in deaf individuals. Face-selective responses in the right dTFA showed a trend for significant positive correlation with face-processing performance in the deaf group [$R_{deaf} = 0.476$, confidence interval (CI) = (−0.101 0.813), P = 0.050] (Fig. 1E). Neither control group showed a similar relationship in the right TVA
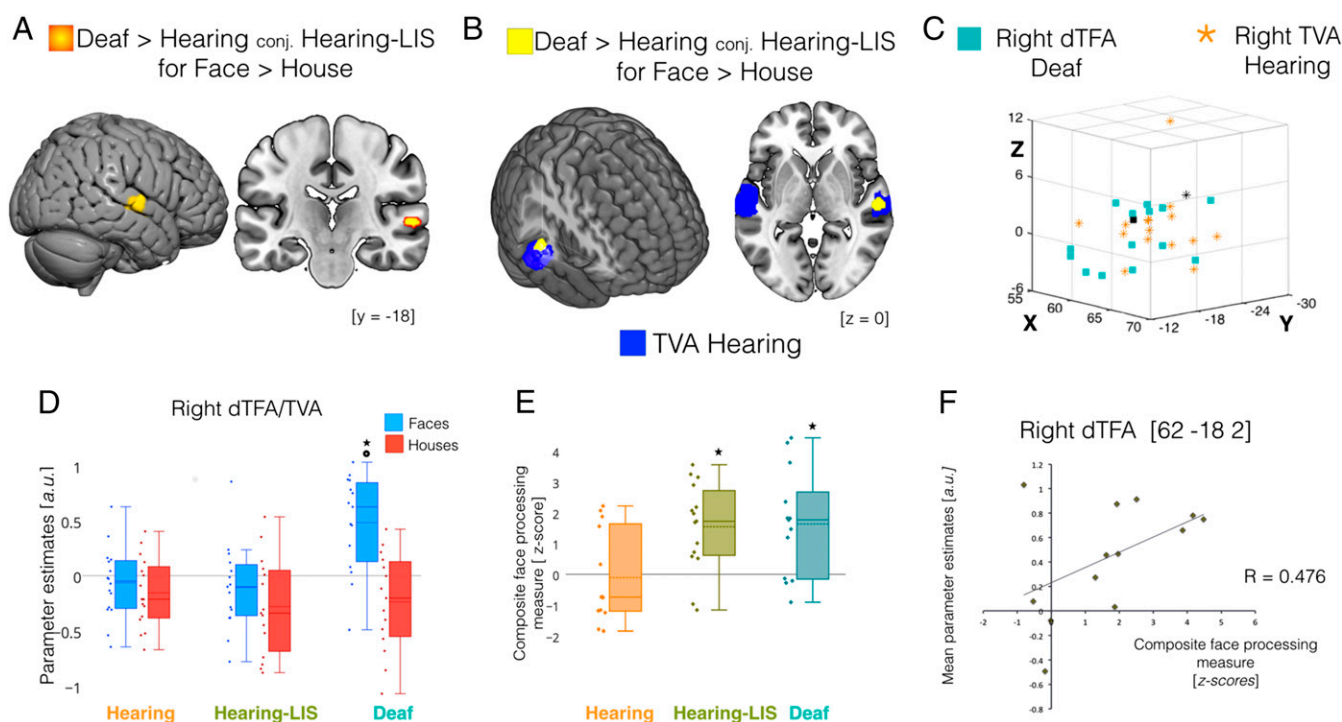

Fig. 1. Cross-modal recruitment of the right dTFA in the deaf. Regional responses significantly differing between groups during face compared with house processing are depicted over multiplanar slices and renders of the MNI-ICBM152 template. (A) Suprathreshold cluster (P < 0.05 FWE small volume-corrected) showing difference between deaf subjects compared with both hearing subjects and hearing LIS users (conj., conjunction analysis). (B) Depiction of the spatial overlap between face-selective response in deaf subjects (yellow) and the voice-selective response in hearing subjects (blue) in the right hemisphere. (C) A 3D scatterplot depicting individual activation peaks in mid STG/STS for face-selective responses in early deaf subjects (cyan squares) and voice-selective responses in hearing subjects (orange stars); black markers represent the group maxima for face selectivity in the right DTFA of deaf subjects (square) and voice selectivity in the right TVA of hearing subjects (star). (D) Box plots showing the central tendency (a.u., arbitrary unit; solid line, median; dashed line, mean) of activity estimates for face (blue) and house (red) processing computed over individual parameters (diamonds) extracted at group maxima for right TVA in each group. *P < 0.001 between groups; P < 0.001 for faces > houses in deaf subjects. (E) Box plots showing central tendency for composite face-processing scores (z-scores; solid line, median; dashed line, mean) for the three groups; *P < 0.016 for deaf > hearing and hearing-LIS > hearing. (F) Scatterplot displaying a trend for significant positive correlation (P = 0.05) between individual face-selective activity estimates and composite measures of face-processing ability in deaf subjects.

[R$_{hearing\ subjects}$ = 0.038, CI = (−0.527 0.57), P = 0.451; R$_{hearing-LIS}$ = −0.053, CI = (−0.55 0.472), P = 0.851]. No significant correlation was detected between neural and behavioral responses to house information deaf subject (R = 0.043, P = 0.884). Moreover, behavioral performances for the house and face tests did not correlate with LIS exposure. It is, however, important to note that the absence of a significant difference in strength of correlation between deaf and hearing groups (see confidence intervals reported above) limits our support for the position that cross-modal reorganization is specifically linked to face perception performance in deaf individuals.

Experiment 2: Reorganized Right dTFA Codes Individual Face Identities. To further evaluate whether reorganized dTFA is also able to differentiate between individual faces, we implemented a second experiment using fMRI adaptation (16). Recent studies in hearing individuals have found that a rapid presentation rate, with a peak at about six face stimuli by second (6 Hz), leads to the largest fMRI-adaptation effect in ventral occipitotemporal face-selective regions, including the FFA, indicating optimal individualization of faces at these frequency rates (21, 22). Participants were presented with blocks of identical or different faces at five frequency rates of presentation between 4 and 8.5 Hz. Individual beta values were estimated for each condition (same/different faces × five frequencies) individually in the right FFA (in all groups), TVA (in hearing subjects and hearing LIS users), and dTFA (in deaf subjects).

Because there were no significant interactions in the TVA [(group) × (identity) × (frequency), P = 0.585] or FFA [(group) × (identity) × (frequency), P = 0.736] or group effects (TVA, P = 0.792; FFA, P = 0.656) when comparing the hearing and hearing-LIS groups, they were merged in a single group for subsequent analyses. With the exception of a main effect of face identity, reflecting the larger response to "different" from identical faces for deaf and hearing participants (Fig. 2B), there were no other significant main or interaction effects in the right FFA. In the TVA/dTFA clusters, in addition to a main effect of face identity (P < 0.001), we also observed two significant interactions of (group) × (face identity) (P = 0.013) and of (group) × (identity) × (frequency) (P = 0.008). A post hoc t test revealed a larger response to different faces (P = 0.034) across all frequencies in deaf compared with hearing participants. In addition, the significant three-way interaction was driven by larger responses to different faces between 4 and 6.6 Hz (4 Hz, P = 0.039; 6 Hz, P = 0.039; 6.6 Hz, P = 0.003) (Fig. 2A) in deaf compared

with hearing participants. In this averaged frequency range, there was a trend for significant release from adaptation in hearing participants (P = 0.031; for this test, the significance threshold was P = 0.05/two groups = 0.025) and a highly significant effect of release in deaf subjects (P < 0.001); when the two groups were directly compared, the deaf group also showed larger release from adaptation compared with hearing and hearing-LIS participants (P < 0.001) (Fig. 2B). These observations not only reveal that the right dTFA shows enhanced coding of individual face identity in deaf individuals but also suggest that the right TVA may show a similar potential in hearing individuals.

Experiment 3: Early Selectivity for Faces in the Right dTFA. In a third neuroimaging experiment, magneto-encephalographic (MEG) responses were recorded during an oddball task with the same face and house images used in the fMRI face localizer. Because no differences were observed between the hearing and hearing-LIS groups for the fMRI face localizer experiment, only deaf subjects (n = 17) and hearing (n = 14) participants were included in this MEG experiment.

Sensor-space analysis on evoked responses to face and house stimuli was performed using permutation statistics and corrected for multiple comparisons with a maximum cluster-mass threshold. Clustering was performed across space (sensors), and time (100 to 300 ms). Robust face-selective responses across groups (P < 0.005, cluster-corrected) were revealed in a large number of sensors mostly around 160 to 210 ms (Fig. 3A), in line with previous observations (23). Subsequent time domain beam forming [linear constrained minimum variance (LCMV)] on this time window of interest showed face-selective regions of the classical face-selective network, including the FFA (Fig. 3 B and C, Top). To test whether dTFA, as identified in fMRI, is already recruited during this early time window of face perception, we tested whether face selectivity was higher in the deaf versus hearing group. For increased statistical sensitivity, a small volume correction was applied using a 15-mm sphere around the voice-selective peak of activation observed in the hearing group in fMRI (MNI x = 63; y = −22; z = 4). Independently reproducing our fMRI results, we observed enhanced selective responses to faces versus houses in deaf compared with hearing subjects, specifically in the right middle temporal gyrus (Fig. 3 B and C, Bottom).

Finally, to explore the timing of face selectivity in dTFA, virtual sensor time courses were extracted for each group and condition from grid points close to the fMRI peak locations showing face selectivity (FFA, hearing and deaf) and voice selectivity (TVA, hearing subjects). We found a face-selective component in dTFA
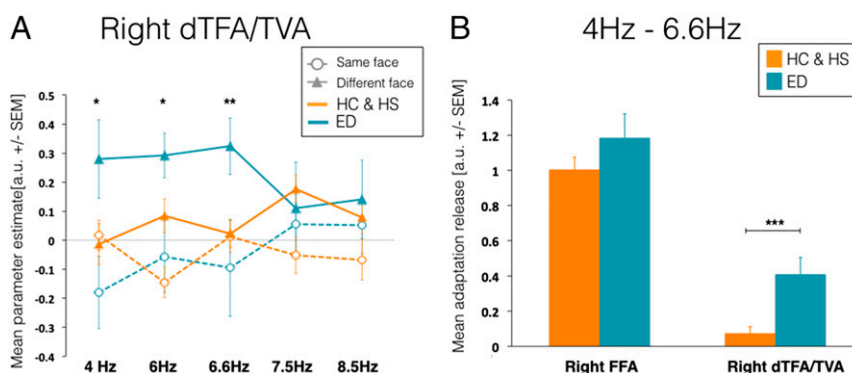
Fig. 2. Adaptation to face identity repetition in the right dTFA of deaf individuals. (A) Mean activity estimates [beta weights; a.u. (arbitrary unit) ± SEM] are reported at each frequency rate of stimulation for same (empty circle/dashed line) and different (full triangle/solid line) faces in both deaf (cyan) and hearing (orange) individuals. Deaf participants showed larger responses for different faces at 4 to 6.6 Hz (*P < 0.05; **P < 0.01) compared with hearing individuals. (B) Bar graphs show the mean adaptation-release estimates (a.u. ± SEM) across frequencies rates 4 to 6.6 Hz in the right FFA and right dTFA/TVA in deaf (orange) and hearing (cyan) individuals. In deaf subjects, the release from adaptation to different faces is above baseline (P < 0.001) and larger than in hearing individuals (***P < 0.001). No significant differences are found in the right FFA. ED, early deaf; HC, hearing controls; HS, hearing-LIS controls.

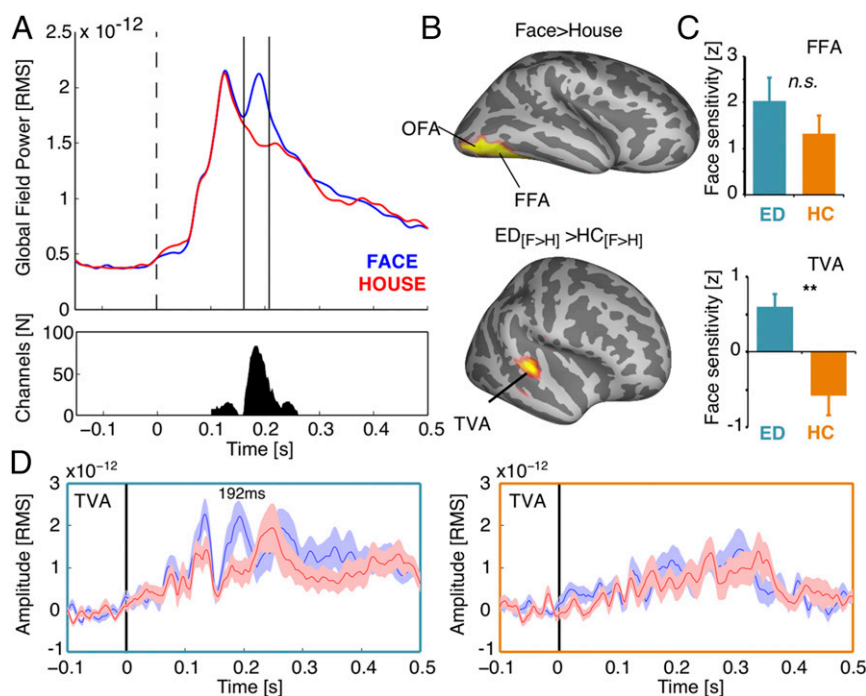**Fig. 3.** Face selectivity in the right dTFA is observed within 200 ms post-stimulus. (*A*) Global field power of the evoked response for faces (blue) and houses (red) across participants. (*Bottom*) The number of sensors contributing to the difference between the two conditions ($P < 0.005$, cluster-corrected) at different points in time. Vertical bars (*Top*) mark the time window of interest (160 to 210 ms) for source reconstruction. (*B, Top*) Face-selective regions within the time window of interest in ventral visual areas across groups ($P < 0.05$, FWE). *Bottom* highlights the interaction effect between groups ($P < 0.05$, FWE). (*C*) Bar graphs illustrate broadband face sensitivity (faces versus houses) for deaf (cyan) and hearing subjects (orange) at peak locations in the FFA and TVA. An interaction effect is observed in dTFA ($P < 0.005$), but not FFA. n.s., not significant; **$P < 0.005$. (*D*) Virtual sensors from the TVA peak locations show the averaged rms time course for faces and houses in the deaf (*Left*, cyan box) and hearing (*Right*, orange box) group. Shading reflects the SEM. Face selectivity in the deaf group peaks at 192 ms in dTFA. No discernible peak is visible in the TVA of the hearing group. ED, early deaf; HC, hearing controls; OFA, occipital face area.

with a peak at 192 ms (Fig. 3*D*), 16 ms after the FFA peak at 176 ms (Fig. 3*D*). In contrast, no difference between conditions was seen at the analogous location in the hearing group (Fig. 3*D*).

Long-Range Connections from V2/V3 Support Face-Selective Response in the Deaf TVA. Previous human and animal studies have suggested that long-range connections with preserved sensory cortices might sustain cross-modal reorganization of sensory-deprived cortices (24). We first addressed this question by identifying candidate areas for the source of cross-modal information in the right dTVA; to this end, a psychophysiological interactions (PPI) analysis was implemented, and the face-selective functional connectivity between the right TVA/dTFA and any other brain regions was explored. During face processing specifically, the right dTFA showed a significant increase of interregional coupling with occipital and fusiform regions in the face-selective network, extending to earlier visual associative areas in the lateral occipital cortex (V2/V3) of deaf individuals only (Fig. 4*A*). Indeed, when face-selective functional connectivity was compared across groups, the effect that differentiated most strongly between deaf and both hearing and hearing-LIS individuals was in the right midlateral occipital gyrus [peak coordinates, x = 42, y = −86, z = 8; z = 5.91, cluster size = 1,561, $P < 0.001$ family-wise error (FWE) cluster- and voxel-corrected] (Fig. 4*A* and *SI Appendix*, Table S5). To further characterize the causal mechanisms and dynamics of the pattern of connectivity observed in the deaf group, we investigated effective connectivity to the right dTFA by combining dynamic causal modeling (DCM) and Bayesian model selection (BMS) in this group. Three different, neurobiologically plausible models were defined based on our observations and previous studies of face-selective effective connectivity in hearing individuals (25): The first

model assumed that a face-selective response in the right dTFA was supported by increased direct "feed-forward" connectivity from early visual occipital regions (right V2/V3); the two alternative models assumed that increased "feedback" connectivity from ventral visual face regions (right FFA) or posterior temporal face regions (right pSTS), respectively, would drive face-selective responses in the right dTFA (Fig. 4*B*). Although the latter two models showed no significant contributions, the first model, including direct connections from the right V2/V3 to right TFA, accounted well for face-selective responses in this region of deaf individuals (exceedance probability = 0.815) (Fig. 4*C*).

Discussion

In this study, we combined state-of-the-art multimodal neuroimaging and psychophysical protocols to unravel how early auditory deprivation triggers specific reorganization of auditory-deprived cortical areas to support the visual processing of faces. In deaf individuals, we report enhanced selective responses to faces in a portion of the mid-STS in the right hemisphere, a region overlapping with the right mid-TVA in hearing individuals (26) that we refer to as the deaf temporal face area. The magnitude of right dTFA recruitment in the deaf subjects showed a trend toward positive correlation with measures of individual face recognition ability in this group. Furthermore, significant increase of neural activity for different faces compared with identical faces supports individual face discrimination in the right dTFA of the deaf subjects. Using MEG, we found that face selectivity in the right dTFA emerges within the first 200 ms after face onset, only slightly later than right FFA activation. Finally, we found that increased long-range connectivity from early visual areas best explained the face-selective response observed in the dTFA of deaf individuals.

Benetti et al.
PNAS | Published online June 26, 2017 | E6441

PSYCHOLOGICAL AND COGNITIVE SCIENCES

Fig. 4. Functional and effective connectivity during face processing in early deaf. (*A*) Psychophysiological interactions (PPI) seeding the right TVA/dTFA. (*Left*) Individual loci of time-series extraction are depicted in red over a cut-out of the right mid STG/STS, showing variability in peak of activation within this region in the deaf group. LF, lateral fissure; m, middle. (*Right*) Suprathreshold (*P* = 0.05 FWE cluster-corrected over the whole brain) face-dependent PPI of right TVA in deaf subjects and significant differences between the deaf and the two control groups are superimposed on the MNI-ICBM152 template. (*B*) The three dynamic causal models (DCMs) used for the study of face-specific effective connectivity in the right hemisphere. Each model equally comprises experimental visual inputs in V3, exogenous connections between regions (gray solid lines), and face-specific modulatory connections to FFA and pSTS (black dashed arrows). The three models differ in terms of the face-specific modulatory connections to dTFA. (*C*) Bayesian model selection showed that a modulatory effect of face from V3 to dTFA best fit the face-selective response observed in the deaf dTFA (*Left*) as depicted in the schematic representation of face-specific information flow (*Right*).

Our findings add to the observation of task-specific cross-modal recruitment of associative auditory regions reported by Bola et al (5): We observed, in early deaf humans, selective cross-modal recruitment of a discrete portion of the auditory cortex for specific and high-level visual processes typically supported by the ventral visual stream in the hearing brain. Additionally, we provide evidence for a functional relationship between recruitment of discrete portions of the auditory cortex and specific perceptual improvements in deaf

individuals. The face-selective cross-modal recruitment of dTFA suggests that cross-modal effects do not occur uniformly across areas of the deaf cortex and supports the notion that cross-modal plasticity is related to the original functional specialization of the colonized brain regions (4, 27). Indeed, temporal voice areas typically involved in an acoustic-based representation of voice identity (28) are shown here to code for facial identity discrimination (Fig. 2*A*), which is in line with previous investigations in

blind humans that have reported that cross-modal recruitment of specific occipital regions by nonvisual inputs follows organizational principles similar to those observed in the sighted. For instance, after early blindness, the lexicographic components of Braille reading elicit specific activations in a left ventral fusiform region that typically responds to visual words in sighted individuals (29) whereas auditory motion selectively activates regions typically selective for visual motion in the sighted (30).

Cross-modal recruitment of a sensory-deprived region might find a "neuronal niche" in a set of circuits that perform functions that are sufficiently close to the ones required by the remaining senses (31). It is, therefore, expected that not all visual functions will be equally amenable to reorganization after auditory deprivation. Accordingly, functions targeting (supramodal) processes that can be shared across sensory systems (32, 33) or benefit from multisensory integration will be the most susceptible to selectively recruit specialized temporal regions deprived of their auditory input (4, 27). Our findings support this hypothesis because the processing of faces and voices shares several common functional features, like inferring the identity, the affective states, the sex, and the age of someone. Along those lines, no selective activity to houses was observed in the temporal cortex of deaf subjects, potentially due to the absence of a common computational ground between audition and vision for this class of stimuli. In hearing individuals, face–voice integration is central to person identity decoding (34), occurs in voice-selective regions (35), and might rely on direct anatomical connections between the voice and face networks in the right hemisphere (15). Our observation of stronger face-selective activations in the right than left mid-STG/STS in deaf individuals further reinforces the notion of functional selectivity in the sensory-deprived cortices. In fact, similarly to face perception in the visual domain, the right midanterior STS regions respond more strongly than the left side to nonlinguistic aspects of voice perception and contribute to the perception of individual identity, gender, age, and emotional state by decoding invariant and dynamic voice features in hearing subjects (34). Moreover, our observation that the right dTFA, similarly to the right FFA, shows fMRI adaptation in response to identical faces suggests that this region is able to process face-identity information. This observation is also comparable with previous findings showing fMRI adaptation to speaker voice identity in the right TVA of hearing individuals (36). In contrast, the observation of face selectivity in the posterior STG for deaf compared with hearing controls, but not hearing-LIS users, supports the hypothesis that regions devoted to speech and multimodal processing in the posterior left temporal cortex might, at least in part, reorganize to process visual aspects of sign language (37).

We know from neurodevelopmental studies that, after an initial period of exuberant synaptic proliferation, projections between the auditory and visual cortices are eliminated either through cell death or retraction of exuberant collaterals during the synaptic pruning phase. The elimination of weaker, unused, or redundant synapses is thought to mediate the specification of functional and modular neuronal networks, such as those supporting face-selective and voice-selective circuitries. However, through pressure to integrate face and voice information for individual recognition (38) and communication (39), phylogenetic and ontogenetic experience may generate privileged links between the two systems, due to shared functional goals. Our findings, together with the evidence of a right dominance for face and voice identification, suggest that such privileged links may be nested in the right hemisphere early during human brain development and be particularly susceptible to functional reorganization after early auditory deprivation. Although overall visual responses were below baseline (deactivation) in the right TVA during visual processing in the hearing groups, a nonsignificant trend for a larger response to faces versus houses (Fig. 1D), as well as a relatively weak face identity adaptation effect, was observed. These results may relate to recent evidence showing both visual unimodal and audiovisual bimodal neuronal subpopulations within early voice-sensitive regions in the right hemisphere of hearing

macaques (35). It is therefore plausible that, in the early absence of acoustic information, the brain reorganizes itself by building on existing cross-modal inputs in the right temporal regions.

The neuronal mechanisms underlying cross-modal plasticity have yet to be elucidated in humans although unmasking of existing synapses, ingrowth of existing connections, and rewiring of new connections are thought to support cortical reorganization (24). Our observation that increased feed-forward effective connectivity from early extrastriate visual regions primarily sustains the face-selective response detected in the right dTFA provides supporting evidence in favor of the view that cross-modal plasticity could occur early in the hierarchy of brain areas and that reorganization of long-range connections between sensory cortices may play a key role in functionally selective cross-modal plasticity. This view is consistent with recent evidence that cross-modal visual recruitment of the pSTS was associated with increased functional connectivity with the calcarine cortex in the deaf although the directionality of the effect was undetermined (40). The hypothesis that the auditory cortex participates in early sensory/perceptual processing after early auditory deprivation, in contrast with previous assumptions that such recruitment manifests only for late and higher level cognitive process (41, 42), also finds support in our MEG finding that a face-selective response occurs at about 196 ms in the right dTFA. The finding that at least 150 ms of information accumulation is necessary for high-level individuation of faces in the cortex (22) suggests that the face-selective response in the right dTFA occurs immediately after the initial perceptual encoding of face identity. Similar to our findings, auditory-driven activity in reorganized visual cortex in congenitally blind individuals was also better explained by direct connections with the primary auditory cortex (43) whereas it depended more on feedback inputs from high-level parietal regions in late-onset blindness (43). The crucial role of developmental periods of auditory deprivation in shaping the reorganization of long-range corticocortical connections remains, however, to be determined.

In summary, these findings confirm that cross-modal inputs might remap selectively onto regions sharing common functional purposes in the auditory domain in early deaf people. Our findings also indicate that reorganization of direct long-range connections between auditory and early visual regions may serve as a prominent neuronal mechanism for functionally selective cross-modal colonization of specific auditory regions in the deaf. These observations are clinically relevant because they might contribute to informing the evaluation of potential compensatory forms of cross-modal plasticity and their role in person information processing after early and prolonged sensory deprivation. Moreover, assessing the presence of such functionally specific cross-modal reorganizations may prove important when considering auditory reafferentation via cochlear implant (1).

Materials and Methods

The research presented in this article was approved by the Scientific Committee of the Centro Interdipartimentale Mente/Cervello (CIMeC) and the Committee for Research Ethics of the University of Trento. Informed consent was obtained from each participant in agreement with the ethical principles for medical research involving human subjects (Declaration of Helsinki; World Medical Association) and the Italian law on individual privacy (D.l. 196/2003).

Participants. Fifteen deaf subjects, 16 hearing subjects, and 15 hearing LIS users participated in the fMRI study. Seventeen deaf and 14 hearing subjects successively participated in the MEG study; because 3 out of 15 deaf participants who were included in the fMRI study could not return to the laboratory and take part in the MEG study, an additional group of 5 deaf participants were recruited for the MEG experiment only. The three groups participating in the fMRI experiment were matched for age, gender, handedness (44), and nonverbal IQ (45) as were the deaf and hearing groups included in the MEG experiment (Tables 2 and 3). No participants had reported neurological or psychiatric history, and all had normal or corrected-to-normal vision. Information on hearing status, history of hearing loss, and use of hearing aids was collected in deaf participants through a structured questionnaire (SI Appendix, Table S1).

Table 2. Demographics, behavioral performances, and Italian Sign Language aspects of the 46 subjects participating in the fMRI experiment

Demographics/cognitive test	Participants in fMRI experiment			Statistics
	Hearing controls (n =16)	Hearing-LIS (n = 15)	Deaf (n = 15)	
Mean age, y (SD)	30.81 (5.19)	34.06 (5.96)	32.26 (7.23)	F-test = 1.079 P value = 0.349
Gender, male/female	8/8	5/10	7/8	χ^2 = 0.97 P value = 0.617
Hand preference, % (right/left ratio)	71.36 (48.12)	61.88 (50.82)	58.63 (52.09)	K–W test = 1.15 P value = 0.564
IQ mean estimate (SD)	122.75 (8.95)	124.76 (5.61)	120.23 (9.71)	F-test = 0.983 P value = 0.384
Composite face recognition z-score (SD)	−0.009 (1.51)**	1.709 (1.40)	1.704 (1.75)	F-test = 5.261 P value = 0.010
LIS exposure, y (SD)	—	25.03 (13.84)	21.35 (9.86)	t test = −0.079 P value = 0.431
LIS acquisition, y (SD)	—	11.42 (8.95)	9.033 (11.91)	M–W U test = 116.5 P value = 0.374
LIS frequency percent time/y (SD)	—	70.69 (44.00)	84.80 (26.09)	M–W U test = 97 P value = 0.441

**P < 0.025 in deaf versus hearing controls and in hearing-LIS versus hearing controls. K–W, Kruskal–Wallis; M–W, Mann–Whitney; SD, standard deviation.

Similarly, information about sign language age of acquisition, duration of exposure, and frequency of use was documented in both the deaf and hearing-LIS group, and no significant differences were observed between the two groups (Tables 2 and 3 and *SI Appendix*, Table S2).

Experimental Design: Behavioral Testing. The long version of the Benton Facial Recognition Test (BFRT) (46) and a delayed face recognition test (DFRT), developed specifically for the present study, were used to obtain a composite measure of individual face identity processing in each group (47). The DFRT was administered 10 to 15 min after completion of the face localizer fMRI experiment and presented the subjects with 20 images for each category (faces and houses), half of which they had previously seen in the scanner (see *Experimental Design: fMRI Face Localizer*). Subjects were instructed to indicate whether they thought they had previously seen the given image.

Experimental Design: fMRI Face Localizer. The face localizer task was administered to the three groups (hearing, hearing-LIS, and deaf) (Table 2). Two categories of stimuli were used: images of faces and houses equated for low-level properties. The face condition consisted of 20 pictures of static faces with neutral expression and in a frontal view (Radboud Faces Database) equally representing male and female individuals (10/10). Similarly, the house condition consisted of 20 full-front photographs of different houses. Low-level image properties (mean luminance, contrast, and spatial frequencies) were equated across stimuli categories by editing them with the SHINE (48) toolbox for Matlab (MathWorks, Inc.). A block-designed one-back identity task was implemented in a single run lasting for about 10 min (*SI Appendix*, Fig. S5). Participants were presented with 10 blocks of 21 s duration for each of the two categories of stimuli. In each block, 20 stimuli of the same condition were presented (1,000 ms, interstimulus interval: 50 ms) on a black background screen; in one to three occasions per block, the exact same stimulus was consecutively repeated that the participant had to detect. Blocks were alternated with a resting baseline condition (cross-fixation) of 7 to 9 s.

Experimental Design: fMRI Voice Localizer and fMRI Face Adaptation. For fMRI voice localizer and fMRI face adaptation experiments, we adapted two fMRI designs previously validated (11, 21). See *SI Appendix, Supporting Information* for a detailed description.

fMRI Acquisition Parameters. For each fMRI experiment, whole-brain images were acquired at the Center for Mind and Brain Sciences (University of Trento) on a 4-Tesla Brucker BioSpin MedSpec head scanner using a standard head coil and gradient echo planar imaging (EPI) sequences. Acquisition parameters for

each experiment are reported in *SI Appendix*, Table S3. Both signing and nonsigning deaf individuals could communicate through overt speech or by using a forced choice button-press code previously agreed with the experimenters. In addition, a 3D MP-RAGE T1-weighted image of the whole brain was also acquired in each participant to provide detailed anatomy [176 slices; echo time (TE), 4.18 ms; repetition time (TR), 2,700 ms; flip angle (FA), 7°; slice thickness, 1 mm].

Behavioral Data Analysis. We computed a composite measure of face recognition with unit-weighted z-scores of the BFRT and DFRT to provide a more stable measure of the underlying face-processing abilities, as well as control for the number of independent comparisons. A detailed description of the composite calculation is reported in *SI Appendix, Supporting Information*.

Functional MRI Data Analysis. We analyzed each fMRI dataset using SPM12 (www.fil.ion.ucl.ac.uk/spm/software/spm12/) and Matlab R2012b (The MathWorks, Inc.). *Preprocessing of fMRI data.* For each subject and for each dataset, the first four images were discarded to allow magnetic saturation effects. The remaining images in each dataset (face localizer, 270; voice localizer, 331; face adaptation, 329; ×3 runs) were visually inspected, and a first manual coregistration between the individual first EPI volume of each dataset, the corresponding magnetization-prepared rapid gradient echo (MP-RAGE) volume, and the T1 Montreal Neurological Institute (MNI) template was performed. Subsequently, in each dataset, the images were corrected for timing differences in slice acquisition, were motion-corrected (six-parameter affine transformation), and were realigned to the mean image of the corresponding sequence. The individual T1 image was segmented into gray and white matter parcellations, and the forward deformation field was computed. Functional EPI images (3-mm isotropic voxels) and the T1 image (1-mm isotropic voxels) were normalized to the MNI space using the forward deformation field parameters, and data were resampled at 2 mm isotropic with a fourth degree B-spline interpolation. Finally, the EPI images in each dataset were spatially smoothed with a Gaussian kernel of 6 mm full width at half maximum (FWHM).

For each fMRI experiment, first-level (single-subject) analysis used a design matrix including separate regressors for the conditions of interest plus realignment parameters to account for residual motion artifacts, as well as outlier regressors; these regressors referred to scans with both large mean displacement and/or weaker or stronger globals. The regressors of interest were defined by convolving boxcars functions representing the onset and offset of stimulation blocks in each experiment by the canonical hemodynamic response function

Table 3. Demographics and behavioral performances of the 31 subjects participating in the MEG experiment

Demographics/cognitive test	Participants in MEG experiment		Statistics
	Hearing controls (n =14)	Deaf (n =17)	
Mean age, y (SD)	30.64 (5.62)	35.47 (8.59)	t test = 1.805 P value = 0.082
Gender male/female	6/8	7/10	χ^2 = 0.009 P value < 0.925
Hand preference, % (right/left ratio)	74.75 (33.45)	78.87 (24.40)	M–W U test = 110 P value = 1
IQ mean estimate (SD)	123.4 (8.41)	117.8 (12.09)	M–W U test = 85.5 P value = 0.567
Benton Face Recognition Task z-score (SD)	−0.539 (0.97)*	0.430 (0.96)	t test = 2.594 P value = 0.016

*P = 0.012 in deaf versus hearing controls.

(HRF). Each design matrix also included a filter at 128 s and auto correlation, which was modeled using an autoregressive matrix of order 1.

fMRI face localizer modeling. Two predictors corresponding to face and house images were modeled, and the contrast (face > house) was computed for each participant; these contrast images were then further spatially smoothed by a 6-mm FWHM before group-level analyses. The individual contrast images of the participants were entered in a one-sample t test to localize regions showing the face-selective response in each group. Statistical inference was made at a corrected cluster level of $P < 0.05$ FWE (with a standard voxel-level threshold of $P < 0.001$ uncorrected) and a minimum cluster-size of 50. Subsequently, a one-way ANOVA was modeled with the three groups as independent factors and a conjunction analysis [deaf$_{(face > house)}$ > hearing$_{(face > house)}$ conjunction with deaf$_{(face > house)}$ > hearing-LIS$_{(face > house)}$] implemented to test for differences between the deaf and the two hearing groups. For this test, statistical inferences were performed also at 0.05 FWE voxel-corrected over a small spherical volume (25-mm radius) located at the peak coordinates of group-specific response to vocal sound in the left and right STG/STS, respectively, in hearing subjects (Table 1). Consequently, measures of individual response to faces and houses were extracted from the right and left TVA in each participant. To account for interindividual variability, a search sphere of 10-mm radius was centered at the peak coordinates (x = 63, y = −22, z = −4; x = −60, y = −16, z = 1; MNI) corresponding to the group maxima for (vocal > nonvocal sounds) in the hearing group. Additionally, the peak-coordinates search was constrained by the TVA masks generated in our hearing group to exclude extraction from posterior STS/STG associative subregions that are known to be also involved in face processing in hearing individuals. Finally, the corresponding beta values were extracted from a 5-mm sphere centered on the selected individual peak coordinates (*SI Appendix, Supporting Information*). These values were then entered in a repeated measure ANOVA with the two visual conditions as within-subject factor and the three groups as between-group factor.

fMRI face-adaptation modeling. We implemented a general linear model (GLM) with 10 regressors corresponding to the (five frequencies × same/different) face images and computed the contrast images for the [same/different face versus baseline (cross-fixation)] test at each frequency rate of visual stimulation. In addition, the contrast image (different versus same faces) across frequency rates of stimulation was also computed in each participant; at the group level, these contrast images were entered as independent variables in three one-sample t tests, separately and specifically for each experimental group, to evaluate whether discrimination of individual faces elicited the expected responses within the face-selective brain network (voxel significance at $P < 0.05$ FWE-corrected). Subsequent analyses were restricted to the functionally defined face- and voice-sensitive areas (voice and face localizers), from which the individual beta values corresponding to each condition were extracted. The Bonferroni correction was applied to correct for multiple comparisons as appropriate.

Region of interest definition for face adaptation. In each participant, ROI definition for face adaptation was achieved by (*i*) centering a sphere volume of 10-mm radius at the peak coordinates reported for the corresponding group, (*ii*) anatomically constraining the search within the relevant cortical gyrus (e.g., for the right FFA, the right fusiform as defined by the Automated Anatomical Labeling Atlas in SPM12), and (*iii*) extracting condition-specific mean beta values from a sphere volume of 5-mm radius (*SI Appendix*, Table S4). The extracted betas were then entered as dependent variables in a series of repeated measures ANOVAs and t tests as reported in *Results*.

Experimental Design: MEG Face Localizer. A face localizer task in the MEG was recorded from 14 hearing (age 30.64) and 17 deaf subjects (age 35.47); all participants, except for 5 deaf subjects, also participated in the fMRI part of the study. Participants viewed the stimulus at a distance from the screen of 100 cm. The images of 40 faces and 40 houses were identical to the ones used in fMRI. After a fixation period (1,000 to 1,500 ms), the visual image was presented for 600 ms. Participants were instructed to press a button whenever an image was presented twice in a row (oddball). Catch trials (~11%) were excluded from subsequent analysis. The images were presented in a pseudorandomized fashion and in three consecutive blocks. Every stimulus was repeated three times, adding up to a total number of 120 trials per condition.

MEG Data Acquisition. MEG was recorded continuously on a 102 triple sensor (two gradiometer, and one magnetometer) whole-head system (Elekta Neuromag). Data were acquired with a sampling rate of 1 kHz and an online band pass filter between 0.1 and 330 Hz. Individual headshapes were recorded using a Polhemus FASTRAK 3D digitizer. The head position was measured continuously using five localization coils (forehead, mastoids). For improved source reconstruction, individual structural MR images were acquired on a 4T scanner (Bruker Biospin).

MEG data analysis.
Preprocessing. The data preprocessing and analysis were performed using the open-source toolbox fieldtrip (49), as well as custom Matlab codes. The continuous data were filtered (high-pass Butterworth filter at 1 Hz; DFT filter at 50, 100, and 150 Hz) and downsampled to 500 Hz to facilitate computational efficiency. Analyses were performed on the gradiometer data. The filtered continuous data were epoched around the events of interest and inspected visually for muscle and jump artifacts. Remaining ocular or cardiac artifacts were removed from the data using extended infomax independent component analysis (ICA), with a weight change stop criterion of 10^{-7}. Finally, a prestimulus baseline of 150 ms was applied to the cleaned epochs.

Sensor-space analysis. Sensor-space analysis was performed across groups before source-space analyses. The cleaned data were low-pass filtered at 30 Hz and averaged separately across face and house trials. Statistical comparisons between the two conditions were performed using a cluster permutation approach in space (sensors) and time (50) in a time window between 100 and 300 ms after stimulus onset. Adjacent points in time and space exceeding a predefined threshold ($P < 0.05$) were grouped into one or multiple clusters, and the summed cluster t values were compared against a permutation distribution. The permutation distribution was generated by randomly reassigning condition membership for each participant (1,000 iterations) and computing the maximum cluster mass on each iteration. This approach reliably controls for multiple comparisons at the cluster level. The time period with the strongest difference between faces and houses was used to guide subsequent source analysis. To illustrate global energy fluctuations during the perception of faces and houses, global field power (GFP) was computed as the root mean square (rms) of the averaged response to the two stimulus types across sensors.

Source-space analysis. Functional data were coregistered with the individual subject MRI using anatomical landmarks (preauricular points and nasion) and the digitized headshape to create a realistic single-shell head model. When no individual structural MRI was available (five participants), a model of the individual anatomy was created by warping an MNI template brain to the individual subject's head shape. Broadband source power was projected onto a 3D grid (8-mm spacing) using linear constrained minimum variance (LCMV) beam forming. To ensure stable und unbiased filter coefficients, a common filter was computed from the average covariance matrix across conditions between 0 and 500 ms after stimulus onset. Whole-brain statistics were performed using a two-step procedure. First, independent-samples t tests were computed for the difference between face and house trials by permuting condition membership (1,000 iterations). The resulting statistical T-maps were converted to Z-maps for subsequent group analysis. Finally, second-level group statistics were performed using statistical nonparametric mapping (SnPM), and family-wise error (FWE) correction at $P < 0.05$ was applied to correct for multiple comparisons. To further explore the time course of face processing in FFA and dTFA for the early deaf participants, virtual sensors were computed on the 40-Hz low-pass filtered data using an LCMV beam former at the FFA and TVA/dTFA locations of interest identified in the whole-brain analysis. Because the polarity of the signal in source space is arbitrary, we computed the absolute for all virtual sensor time series. A baseline correction of 150 ms prestimulus was applied to the data.

fMRI Functional Connectivity Analysis. Task-dependent contributions of the right dTFA and TVA to brain face-selective responses elsewhere were assessed in the deaf and in the hearing groups, respectively, by implementing a psychophysiological interactions (PPI) analysis (51) on the fMRI face localizer dataset. The individual time series for the right TVA/dTFA were obtained by extracting the first principal component from all raw voxel time series in a sphere (radius of 5 mm) centered on the peak coordinates of the subject-specific activation in this region (i.e., face-selective responses in deaf subjects and voice-selective responses in hearing subjects and hearing LIS users). After individual time series had been mean-corrected and high-pass filtered to remove low-frequency signal drifts, a PPI term was computed as the element-by-element product of the TVA/dTFA time series and a vector coding for the main effect of task (1 for face presentation, −1 for house presentation). Subsequently, a GLM was implemented including the PPI term, region-specific time series, main effect of task vector, moving parameters, and outlier scans vector as model regressors. The contrast image corresponding to the positive-tailed one-sample t test over the PPI regressor was computed to isolate brain regions receiving stronger contextual influences from the right TVA/dTFA during face processing compared with house processing. The opposite (i.e., stronger) influences during house processing were achieved by computing the contrast image for the negative-tailed one-sample t test over the same PPI regressor. These subject-specific contrast images were spatially smoothed by a 6-mm FWHM prior submission to subsequent statistical analyses. For each group, individually smoothed contrast images were entered as the dependent variable in a one-sample t test to isolate regions showing face-specific increased functional connectivity with the right TVA/dTFA.

Finally, individual contrast images were also entered as the dependent variable in two one-way ANOVAs, one for face and one for house responses, with the three groups as between-subject factor to detect differences in functional connectivity from TVA/dTFA between groups. For each test, statistical inferences were made at corrected cluster level of $P < 0.05$ FWE (with a standard voxel-level threshold of $P < 0.001$ uncorrected) with a minimum size of 50 voxels.

Effective Connectivity Analysis. Dynamic causal modeling (DCM) (52), a hypothesis-driven analytical approach, was used to characterize the causality between the activity recorded in the set of regions that showed increased functional connectivity with the right dTFA in the deaf group during face compared with house processing. To this purpose, our model space was operationalized based on three neurobiologically plausible and sufficient alternatives: (*i*) Face-selective response in the right dTFA is supported by increased connectivity modulation directly from the right V2/V3, (*ii*) face-selective response in the right dTFA is supported indirectly by increased connectivity modulation from the right FFA, or (*iii*) face-selective response in the right dTFA is supported indirectly by increased connectivity modulation from the right pSTS. DCM models can be used for investigating only brain

responses that present a relation to the experimental design and can be observed in each individual included in the investigation (52). Because no temporal activation was detected for face and house processing in hearing subjects and hearing LIS users, these groups were not included in the DCM analysis. For a detailed description of DCMs, see *SI Appendix, Supporting Information*.

The three DCMs were fitted with the data from each of the 15 deaf participants, which resulted in 45 fitted DCMs and corresponding log-evidence and posterior parameters estimates. Subsequently, random-effect Bayesian model selection (53) was applied to the estimated evidence for each model to compute the "exceedance probability," which is the probability of each specific model to better explain the observed activations compared with any other model.

1. Heimler B, Weisz N, Collignon O (2014) Revisiting the adaptive and maladaptive effects of crossmodal plasticity. *Neuroscience* 283:44–63.
2. Finney EM, Fine I, Dobkins KR (2001) Visual stimuli activate auditory cortex in the deaf. *Nat Neurosci* 4:1171–1173.
3. Karns CM, Dow MW, Neville HJ (2012) Altered cross-modal processing in the primary auditory cortex of congenitally deaf adults: A visual-somatosensory fMRI study with a double-flash illusion. *J Neurosci* 32:9626–9638.
4. Lomber SG, Meredith MA, Kral A (2010) Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nat Neurosci* 13:1421–1427.
5. Bola Ł, et al. (2017) Task-specific reorganization of the auditory cortex in deaf humans. *Proc Natl Acad Sci USA* 114:E600–E609.
6. Pavani F, Bottari D (2012) Visual abilities in individuals with profound deafness: A critical review. *The Neural Bases of Multisensory Processes*, eds Murray MM, Wallace MT (CRC Press, Boca Raton, FL).
7. Lee H-J, et al. (2007) Cortical activity at rest predicts cochlear implantation outcome. *Cereb Cortex* 17:909–917.
8. Bettger J, Emmorey K, McCullough S, Bellugi U (1997) Enhanced facial discrimination: Effects of experience with American sign language. *J Deaf Stud Deaf Educ* 2:223–233.
9. Rouger J, Lagleyre S, Démonet JF, Barone P (2012) Evolution of crossmodal reorganization of the voice area in cochlear-implanted deaf patients. *Hum Brain Mapp* 33:1929–1940.
10. Stropahl M, et al. (2015) Cross-modal reorganization in cochlear implant users: Auditory cortex contributes to visual face processing. *Neuroimage* 121:159–170.
11. Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. *Nature* 403:309–312.
12. Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci* 25:5004–5012.
13. Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311.
14. von Kriegstein K, Kleinschmidt A, Sterzer P, Giraud A-L (2005) Interaction of face and voice areas during speaker recognition. *J Cogn Neurosci* 17:367–376.
15. Blank H, Anwander A, von Kriegstein K (2011) Direct structural connections between voice- and face-recognition areas. *J Neurosci* 31:12906–12915.
16. Grill-Spector K, Malach R (2001) fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107:293–321.
17. Jacques C, et al. (2016) Corresponding ECoG and fMRI category-selective signals in human ventral temporal cortex. *Neuropsychologia* 83:14–28.
18. Benton AL, Van Allen MW (1968) Impairment in facial recognition in patients with cerebral disease. *Trans Am Neurol Assoc* 93:38–42.
19. Hildebrandt A, Sommer W, Herzmann G, Wilhelm O (2010) Structural invariance and age-related performance differences in face cognition. *Psychol Aging* 25:794–810.
20. Arnold P, Murray C (1998) Memory for faces and objects by deaf and hearing signers and hearing nonsigners. *J Psycholinguist Res* 27:481–497.
21. Gentile F, Rossion B (2014) Temporal frequency tuning of cortical face-sensitive areas for individual face perception. *Neuroimage* 90:256–265.
22. Alonso-Prieto E, Belle GV, Liu-Shuang J, Norcia AM, Rossion B (2013) The 6 Hz fundamental stimulation frequency rate for individual face discrimination in the right occipito-temporal cortex. *Neuropsychologia* 51:2863–2875.
23. Halgren E, Raij T, Marinkovic K, Jousmäki V, Hari R (2000) Cognitive response profile of the human fusiform face area as determined by MEG. *Cereb Cortex* 10:69–81.
24. Bavelier D, Neville HJ (2002) Cross-modal plasticity: Where and how? *Nat Rev Neurosci* 3:443–452.
25. Lohse M, et al. (2016) Effective connectivity from early visual cortex to posterior occipitotemporal face areas supports face selectivity and predicts developmental prosopagnosia. *J Neurosci* 36:3821–3828.
26. Pernet CR, et al. (2015) The human voice areas: Spatial organization and inter-individual variability in temporal and extra-temporal cortices. *Neuroimage* 119:164–174.
27. Dormal G, Collignon O (2011) Functional selectivity in sensory-deprived cortices. *J Neurophysiol* 105:2627–30.
28. Latinus M, McAleer P, Bestelmeyer PEG, Belin P (2013) Norm-based coding of voice identity in human auditory cortex. *Curr Biol* 23:1075–1080.
29. Reich L, Szwed M, Cohen L, Amedi A (2011) A ventral visual stream reading center independent of visual experience. *Curr Biol* 21:363–368.
30. Collignon O, et al. (2011) Functional specialization for auditory-spatial processing in the occipital cortex of congenitally blind humans. *Proc Natl Acad Sci USA* 108:4435–4440.
31. Collignon O, Voss P, Lassonde M, Lepore F (2009) Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. *Exp Brain Res* 192:343–358.
32. Pascual-Leone A, Hamilton R (2001) The metamodal organization of the brain. *Prog Brain Res* 134:427–445.
33. Amedi A, Malach R, Hendler T, Peled S, Zohary E (2001) Visuo-haptic object-related activation in the ventral visual pathway. *Nat Neurosci* 4:324–330.
34. Yovel G, Belin P (2013) A unified coding strategy for processing faces and voices. *Trends Cogn Sci* 17:263–271.
35. Perrodin C, Kayser C, Logothetis NK, Petkov CI (2014) Auditory and visual modulation of temporal lobe neurons in voice-sensitive and association cortices. *J Neurosci* 34:2524–2537.
36. Belin P, Zatorre RJ (2003) Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport* 14:2105–2109.
37. MacSweeney M, Capek CM, Campbell R, Woll B (2008) The signing brain: The neurobiology of sign language. *Trends Cogn Sci* 12:432–440.
38. Sheehan MJ, Nachman MW (2014) Morphological and population genomic evidence that human faces have evolved to signal individual identity. *Nat Commun* 5:4800.
39. Ghazanfar AA, Logothetis NK (2003) Neuroperception: Facial expressions linked to monkey calls. *Nature* 423:937–938.
40. Shiell MM, Champoux F, Zatorre RJ (2015) Reorganization of auditory cortex in early-deaf people: Functional connectivity and relationship to hearing aid use. *J Cogn Neurosci* 27:150–163.
41. Leonard MK, et al. (2012) Signed words in the congenitally deaf evoke typical late lexicosemantic responses with no early visual responses in left superior temporal cortex. *J Neurosci* 32:9700–9705.
42. Ding H, et al. (2015) Cross-modal activation of auditory regions during visuo-spatial working memory in early deafness. *Brain* 138:2750–2765.
43. Collignon O, et al. (2013) Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. *Brain* 136:2769–2783.
44. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9:97–113.
45. Raven J, Raven JC, Court J (1998) *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Available at https://books.google.es/books?id=YrvAAQAACAAJ&hl=es. Accessed June 15, 2017.
46. Benton A, Hamsher K, Varney N, Spreen O (1994) *Contribution to Neuropsychological Assessment* (Oxford Univ Press, New York).
47. Ackerman PL, Cianciolo AT (2000) Cognitive, perceptual-speed, and psychomotor determinants of individual differences during skill acquisition. *J Exp Psychol Appl* 6:259–290.
48. Willenbockel V, et al. (2010) Controlling low-level image properties: The SHINE toolbox. *Behav Res Methods* 42:671–684.
49. Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
50. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.
51. Friston KJ, et al. (1997) Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6:218–229.
52. Stephan KE, et al. (2007) Dynamic causal models of neural system dynamics: Current state and future extensions. *J Biosci* 32:129–144.
53. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. *Neuroimage* 46:1004–1017.