Neuronal populations in the occipital cortex of the blind synchronize to the temporal dynamics of speech

Markus Johannes Van Ackeren[1], Francesca M Barbero[2,3], Stefania Mattioni[1], Roberto Bottini[1], Olivier Collignon[1,2,3]*

[1]Center for Mind/Brain Studies, University of Trento, Trento, Italy; [2]Institute of research in Psychology, University of Louvain, Louvain, Belgium; [3]Institute of Neuroscience, University of Louvain, Louvain, Belgium

Abstract The occipital cortex of early blind individuals (EB) activates during speech processing, challenging the notion of a hard-wired neurobiology of language. But, at what stage of speech processing do occipital regions participate in EB? Here we demonstrate that parieto-occipital regions in EB enhance their synchronization to acoustic fluctuations in human speech in the theta-range (corresponding to syllabic rate), irrespective of speech intelligibility. Crucially, enhanced synchronization to the intelligibility of speech was selectively observed in primary visual cortex in EB, suggesting that this region is at the interface between speech perception and comprehension. Moreover, EB showed overall enhanced functional connectivity between temporal and occipital cortices that are sensitive to speech intelligibility and altered directionality when compared to the sighted group. These findings suggest that the occipital cortex of the blind adopts an architecture that allows the tracking of speech material, and therefore does not fully abstract from the reorganized sensory inputs it receives.
DOI: https://doi.org/10.7554/eLife.31640.001

Introduction

The human cortex comprises a number of specialized units, functionally tuned to specific types of information. How this functional architecture emerges, persists, and develops throughout a person's life are among the most challenging and exciting questions in neuroscience research. Although there is little debate that both genetic and environmental influences affect brain development, it is currently not known how these two factors shape the functional architecture of the cortex. A key topic in this debate is the organization of the human language system. Language is commonly thought to engage a well-known network of regions around the lateral sulcus. The consistency of this functional mapping across individuals and its presence early in development are remarkable, and often used to argue that the neurobiological organization of the human language system is the result of innate constraints (*Dehaene-Lambertz et al., 2006*; *Berwick et al., 2013*). Does the existence of a highly consistent set of regions for language acquisition and processing imply that this network is 'hard-wired' and immutable to experience? Strong nativist theories for linguistic innateness leave little room for plasticity due to experience (*Bates, 1999*), suggesting that we should conceive 'the growth of language as analogous to the development of a bodily organ' (*Chomsky, 1976*, p.11). However, studies in infants born with extensive damage to cortical regions that are typically involved in language processing may develop normal language abilities, thereby demonstrating that the language network is subject to reorganization (*Bates, 2005*). Perhaps the most intriguing demonstrations to show that the neurobiology of language is susceptible to change due to experiencecome from studies showing functional selectivity to language in primary and secondary 'visual' areas in

congenitally blind individuals (*Röder et al., 2002*; *Burton, 2003*; *Amedi et al., 2004*; *Bedny et al., 2011*; *Arnaud et al., 2013*). Such reorganization of the language network is particularly fascinating because it arises in the absence of injury to the core language network (*Bates, 2005*; *Atilgan et al., 2017*).

However, the level at which the occipital cortex is involved in speech representation in the early blind (EB), remains poorly understood. Speech comprehension requires that the brain extracts meaning from the acoustic features of sounds (*de Heer et al., 2017*). Although several neuroimaging studies have yielded valuable insights about the processing of speech in EB adults (*Arnaud et al., 2013*; *Bedny et al., 2011*; *Büchel, 2003*; *Lane et al., 2015*; *Röder et al., 2002*) and infants (*Bedny et al., 2015*), these methods do not adequately capture the fast and continuous nature of speech processing. Because speech unfolds over time, understanding spoken language relies on the ability to track the incoming acoustic signal in near real-time (*Peelle and Davis, 2012*). Indeed, speech is a fluctuating acoustic signal that rhythmically excites neuronal populations in the brain (*Poeppel et al., 2008*; *Gross et al., 2013*; *Peelle et al., 2013*). Several studies have demonstrated that neuronal populations in auditory areas entrain to the acoustic fluctuations that are present in human speech around the syllabic rate (*Luo and Poeppel, 2007*; *Kayser et al., 2009*; *Szymanski et al., 2011*; *Zoefel and VanRullen, 2015*). It has therefore been suggested that entrainment reflects a key mechanism underlying hearing by facilitating the parsing of individual syllables through adjusting the sensory gain relative to fluctuations in the acoustic energy (*Giraud and Poeppel, 2012*; *Peelle and Davis, 2012*; *Ding and Simon, 2014*). Crucially, because some regions that track the specific acoustic rhythm of speech are sensitive to speech intelligibility, neural synchronization is not only driven by changes in the acoustic cue of the auditory stimuli, but also reflects cortical encoding and processing of the auditory signal (*Peelle et al., 2013*; *Ding and Simon, 2014*). Speech tracking is therefore an invaluable tool to probe regions that interface speech perception and comprehension (*Poeppel et al., 2008*; *Gross et al., 2013*; *Peelle et al., 2013*).

Does the occipital cortex of EB people synchronize to speech rhythm? Is this putative synchronization of neural activity to speech influenced by comprehension? Addressing these questions would provide novel insights into the functional organization of speech processing rooted in the occipital cortex of EB people. In the current study, we investigated whether neuronal populations in blind occipital cortex synchronize to rhythmic dynamics of speech, by relating the amplitude fluctuations in speech to electromagnetic dynamics recorded from the participant's brain. To this end, we quantified the local brain activity and directed connectivity in a group of early blind (EB; n = 17) and sighted individuals (SI; n = 16) using magnetoencephalography (MEG) while participants listened to short narrations from audiobooks. If the occipital cortex of the blind entrains to speech rhythms, this will support the idea that this region processes low-level acoustic features relevant for understanding language. We further tested whether the putative synchronization of occipital responses in EB people benefits from linguistic information or only relates to acoustic information. To separate linguistic and acoustic processes, we relied on a noise-vocoding manipulation. This method spectrally distorted the speech signal in order to impair intelligibility gradually, but systematically preserves the slow amplitude fluctuations responsible for speech rhythm (*Shannon et al., 1995*; *Peelle and Davis, 2012*). Furthermore, to go beyond differences in the local encoding of speech rhythm in the occipital cortex, we also investigated whether the connectivity between occipital and temporal regions sensitive to speech comprehension is altered in early blindness.

Results

Story comprehension

Participants listened to either natural speech segments (nat-condition) or to altered vocoded versions of these segments (see Materials and methods for details). In the 8-channel vocoded condition, the voice of the speaker is highly distorted but the intelligibility is unperturbed. By contrast, in the 1-channel vocoded condition, the speech is entirely unintelligible. After listening to each speech segment, participants were provided with a short statement about the segment, and asked to indicate whether the statement was true or false. Behavioural performance on these comprehension

statements was analysed using linear mixed-effects models with maximum likelihood estimation. This method is a linear regression that takes into account dependencies in the data, as present in repeated measures designs. Blindness and intelligibility were included as fixed effects, while subject was modelled as a random effect. Intelligibility was nested in subjects. Intelligibility had a significant effect on story comprehension (χ(2)=110.7, p<0.001). The effect of blindness and the interaction between intelligibility and blindness were non-significant (χ(1)=1.14, p=0.286 and χ(2)=0.24, p=0.889). Orthogonal contrasts demonstrated that speech comprehension was stronger in the nat and 8-channel condition versus the 1-channel condition (b = 0.78, t(62)=14.43, p<0.001, r = 0.88). There was no difference between the nat and the 8-channel condition (b = 0.02, t(62)=1.39, p=0.17). Thus, speech intelligibility was reduced in the 1-channel, but not the 8-channel vocoded condition (*Figure 1D*). The lack of effect for the factor blindness suggests that there is no evidence for a potential difference in comprehension, attention or motivation between groups. Descriptive statistics of the group, and condition means are depicted in *Table 1*.

Tracking intelligible speech in blind and sighted temporal cortex

The overall coherence spectrum, which highlights the relationship between the amplitude envelope of speech and the signal recorded from the brain, was maximal over temporal sensors between 6 Hz and 7 Hz (*Figure 1E*). This first analysis was performed on the combined dataset, and hence is not prone to bias or circularity for subsequent analyses, targeting group differences. The peak in the current study is slightly higher than those reported in previous studies (*Gross et al., 2013*). A likely explanation for this shift is the difference in syllabic rate between English (~6.2 Hz), used in previous studies, and Italian (~7 Hz) (*Pellegrino et al., 2011*). The syllabic rate is the main carrier of amplitude fluctuations in speech, and thus is most prone to reset oscillatory activity.

To capture the temporal scale of cerebro-acoustic coherence effects (6–7 Hz) optimally, and to achieve a robust source estimate, source reconstruction was performed on two separate frequency windows (6 ± 2, and 7 ± 2 Hz) for each subject and condition. The two source images were averaged for subsequent analysis, yielding a single image representing the frequency range 4–9 Hz. By combining the two frequency-bands, we acquire a source estimate that emphasizes the center of our frequency band of interest (6–7 Hz) and tapers off towards the edges. This source estimate optimally represents the coherence spectrum observed in sensor space. The choice of the center frequency was also restricted by the length of the time window used for the analysis. That is, with a 1 s time window and a resulting 1 Hz frequency resolution, a non-integer center frequency at, for example, 6.5 Hz was not feasible. The source-reconstructed coherence at the frequency-range of interest confirmed that cerebro-acoustic coherence was strongest across groups and conditions in bilateral temporal lobes, including primary auditory cortex. (*Figure 1F*).

To test whether envelope tracking in the current study is modulated by intelligibility, we compared the coherence maps for the intelligible (nat) versus non-intelligible (1-channel) condition, in all groups combined (i.e., EB and SI), with dependent-samples permutation t-tests in SPM. The resulting statistical map (*Figure 1G*, p<0.05, FWE-corrected) revealed a cluster in right superior and middle temporal cortex (STG, MTG), where synchronization to the envelope of speech was stronger when participants had access to the content of the story. In other words, the effect of intelligibility on the sensory response (coherence with the speech envelope) suggests an interface between acoustic speech processing and comprehension in STG that is present in both groups.

Speech tracking in blind parieto-occipital cortex

Whether EB individuals recruit additional neural substrate for tracking the envelope of speech was tested using independent-samples permutation t-tests in SPM, contrasting coherence maps between EB and SI for all three conditions combined. The statistical maps (*Figure 1H*, p<0.05, FWE-corrected) revealed enhanced coherence in EB versus SI in parieto-occipital cortex along the medial wall, centered on bilateral Precuneus, branching more extensively into the right hemisphere (see *Table 2*). This main effect of group highlights that neuronal populations in blind parieto-occipital cortex show enhanced synchronization to the acoustic speech signal. That is, blind participants show a stronger sensitivity to the sensory properties of the external speech stimulus at this level.

Finally, to investigate whether and where envelope tracking is sensitive to top-down predictions during intelligible speech comprehension, we subtracted the unintelligible (1-channel) from the
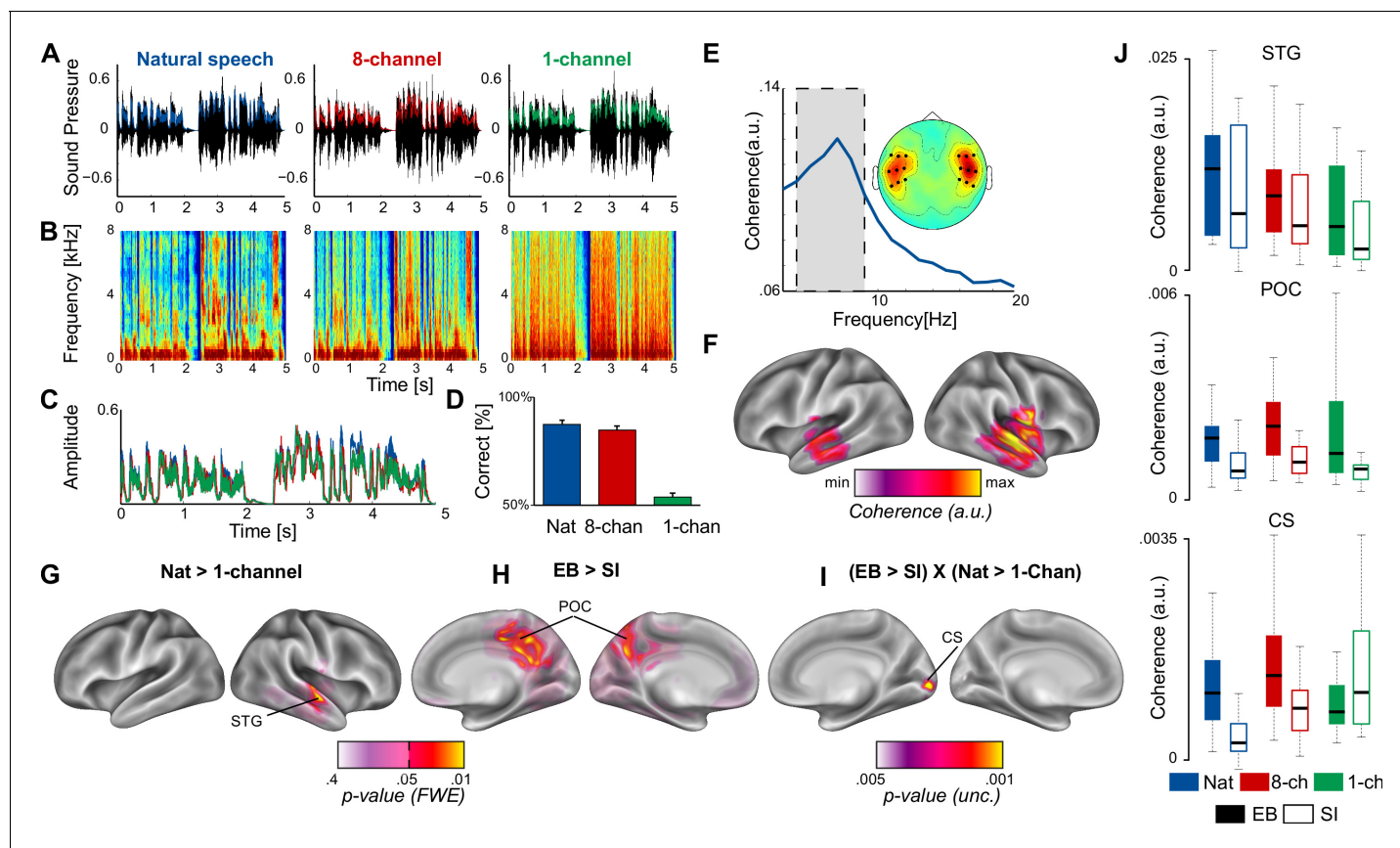
Figure 1. Cerebro-acoustic coherence in EB and SI. (**A**) Segments from each condition (nat, 8-channel, and 1-channel) in the time domain (see also **Video 1**). Overlaid is the amplitude envelope of each condition. (**B**) Spectrogram of the speech samples from (**A**). The effect of spectral vocoding on the fine spectral detail in all three conditions. (**C**) The superimposed amplitude envelopes of the samples for each condition are highly similar, despite distortions in the fine spectral detail. (**D**) Behavioural performance on the comprehension statements reveals that comprehension is unperturbed in the nat and 8-channel conditions, whereas the 1-channel condition elicits chance performance. (**E**) Coherence spectrogram extracted from bilateral temporal sensors reveals a peak in cerebro-acoustic coherence at 7 Hz, across groups and conditions. The shaded area depicts the frequency range of interest in the current study. The topography shows the spatial extent of coherence values in the range 4–9 Hz. Enhanced coherence is observed selectively in bilateral temporal sensors. (**F**) Source reconstruction of the raw coherence confirms that cerebro-acoustic coherence is strongest in the vicinity of auditory cortex, bilaterally, extending into superior and middle temporal gyrus. (**G**) Statistically thresholded maps for the contrast between natural and 1-channel vocoded speech show an effect of intelligibility (p<0.05, FWE-corrected) in right STG. (**H**) Enhanced envelope tracking is observed for EB versus SI in a bilateral parieto-occipital network along the medial wall centred on Precuneus (p<0.05, FWE-corrected). (**I**) The statistical map shows the interaction effect between blindness and intelligibility: Early blind (EB) individuals show enhanced synchronization during intelligible (nat) versus non-intelligible speech (1-channel) as compared to SI in right calcarine sulcus (CS) (p<0.005, uncorrected). (**J**) Boxplots for three regions identified in the whole-brain analysis (top panel, STG; middle panel, parieto-occipital cortex; bottom panel, calcarine sulcus).

intelligible (nat) condition, and computed independent-samples t-tests between groups. As highlighted in the introduction, we were particularly interested in the role played by the Calcarine sulcus (V1) in processing semantic attributes of speech (**Burton et al., 2002**; **Röder et al., 2002**; **Amedi et al., 2003**), and therefore we restricted the statistical analysis to the area around the primary visual cortex. The search volume was constructed from four 10 mm spheres around coordinates in bilateral Calcarine sulcus ([−7,–81, −3]; [−6,–85, 4]; [12, -87, 0; 10, –84, 6]). The coordinates were extracted from a previous study on speech comprehension in EB (**Burton et al., 2002**). The resulting mask also include regions highlighted in similar studies by other groups (**Röder et al., 2002**; **Bedny et al., 2011**).

A significant effect of the interaction between blindness and intelligibility [(nat$_{EB}$ − 1-channel$_{EB}$) − (nat$_{SI}$ − 1-channel$_{SI}$)] was observed in right calcarine sulcus. To explore the spatial specificity of the

Table 1. Proportion of correct responses for the story comprehension questions.
Depicted are the number of subjects (N) in the early blind (EB) and sighted individuals (SI) groups, the mean (M), and the standard error of the mean (SE) for each of the three conditions (Natural, 8-channel, 1-channel).
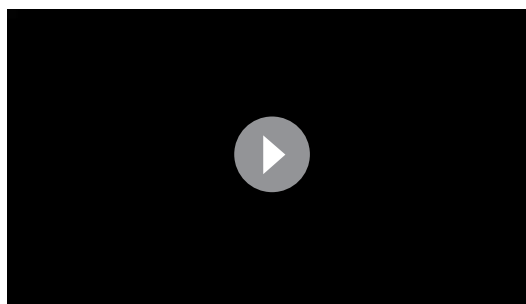
		Natural		8-channel		1-channel	
Group	N	M	SE	M	SE	M	SE
EB	17	0.899	0.019	0.853	0.031	0.559	0.027
SI	15	0.914	0.020	0.890	0.025	0.576	0.034

DOI: https://doi.org/10.7554/eLife.31640.002

effect, we show the whole-brain statistical map for the interaction between intelligibility and blindness at a more liberal threshold (p<0.005, uncorrected) in *Figure 1I*. This shows that intelligible speech selectively engages the area around the right calcarine sulcus in the blind versus the sighted. Specifically, the region corresponding to right 'visual' cortex showed enhanced sensitivity to intelligible speech in EB versus SI (p<0.05, FWE-corrected). However, the analysis contrasting the two intelligible conditions (nat and 8-channel) did not yield a significant effect in the EB, suggesting that the low-level degrading of the stimulus alone does not drive the effect, but rather the intelligibility of the speech segment. Follow-up post-hoc comparisons between the two groups revealed that coherence was stronger during the intelligible condition for EB versus SI (t[30] = 3.09, p=0.004), but not during the unintelligible condition (t[30] = −1.08, p=0.29).

A group difference between EB and SI across conditions was not observed in the calcarine region (t[30] = 1.1, p=0.28). The lack of an overall group effect in calcarine sulcus suggests that there is not a simple enhanced sensory response to speech in the blind. This was different to the response we observed in parietal cortex, in which synchronization was stronger in the blind than in the sighted. Rather, the response in calcarine sulcus only differed between the two groups when speech was intelligible. While overall coherence with the speech envelope was equally high in EB and SI, the blind population showed significantly higher coherence in the intelligible speech condition than did the sighted, who showed higher cerebro-acoustic coherence in the unintelligible condition (1-Chan). The latter is reminiscent of the fact that the more adverse the listening condition (low signal-to-noise-ratio or audio-visual incongruence), the more the visual cortex is entrained to the visual speech signal of actual acoustic speech when presented together with varying levels of acoustic noise (*Park et al., 2016*; *Giordano et al., 2017*). Moreover, *Giordano et al. (2017)* showed an increase of directed connectivity between superior frontal regions and visual cortex under the most challenging (acoustic noise and uninformative visual cues) conditions, again suggesting a link between the reorganization observed in the occipital cortex of blind individuals and typical multisensory pathways involving the occipital cortex in audio-visual speech processing (*Kayser et al., 2008*). Visual deprivation since birth, however, triggers a functional reorganization of the calcarine region that can then dynamically interact with the intelligibility of the speech signal. For illustration purposes, cerebro-acoustic coherence from functional peak locations for intelligibility (across groups) in STG, blindness in the parieto-occipital cortex (*Figure 1G*; POC), and the interaction are represented as boxplots in *Figure 1J*.

Changes in occipito-temporal connectivity in the EB

To further investigate whether cerebro-acoustic peaks in CS and STG which are sensitive to intelligible speech in EB are indicative of a more general re-organization of the network, we conducted functional connectivity analysis. Statistical analysis of the connectivity estimates was

Video 1. Movie of the natural and two vocoding conditions used for one exemplar of auditory segment used in our experiment.
DOI: https://doi.org/10.7554/eLife.31640.006

Table 2. MNI (Montreal Neurological Institute) coordinates and p-values for the three contrasts tested with threshold-free cluster enhancement (TFCE).

		MNI-coordinates (mm)		
	p-value (FWE-cor)	x	y	z
Intelligibility				
Superior temporal gyrus	0.027	64	−16	0
Blindness				
Posterior cingulate	0.030	8	−40	40
Postcentral sulcus	0.033	16	−32	64
Precuneus	0.033	-8	−64	48
Blindness x intelligibility				
Calcarine sulcus	0.007	16	−88	0

DOI: https://doi.org/10.7554/eLife.31640.004

performed on the mean phase-locking value in the theta (4–8 Hz) range.

Using linear mixed-effects models, blindness (EB, SI), intelligibility (nat, 1-channel), and the inter-action between blindness and intelligibility were added to the model in a hierarchical fashion. Blind-ness and intelligibility were modelled as fixed effects, while subject was a random effect. Intelligibility was nested in subject. A main effect was observed only for blindness ($\chi[1] = 4.32$, p=0.038) and was caused by greater connectivity for EB versus SI (*Figure 2A–B*). The main effects of intelligibility and of the interaction between intelligibility and blindnesswere non-significant (p=0.241 and p=0.716, respectively).

Subsequently, linear mixed-effects models were applied to test for the effects of blindness and intelligibility on the directional information flow (PSI) between CS and STG. The fixed and random effects structure was the same as that described in the previous analysis. Here, only the main effect of blindness was significant ($\chi[1] = 4.54$, p=0.033), indicating that the directional information flow differs between groups. The effect of intelligibility and the interaction between intelligibility and blindness were both non-significant (p=0.51 and p=0.377, respectively). Follow-up post-hoc one-sample t-tests on the phase slope estimates for each group individually revealed a significant direc-tion bias from CS to STG for SI ($t[14] = -2.22$, p=0.044). No directional bias was found for EB ($t[16] = 0.74$, p=0.47). As depicted in *Figure 2C–D*, these results suggest that CS projects predomi-nantly to STG in SI, whereas in EB, this interaction is more balanced and trending in the opposite direction.

Discussion

To test whether neuronal populations in the early blind's occipital cortex synchronize to the temporal dynamics present in human speech, we took advantage of the statistical and conceptual power offered by correlating magnetoencephalographic recordings with the envelope of naturalistic contin-uous speech. Using source-localized MEG activity, we confirm and extend previous research (*Scott et al., 2000*; *Davis and Johnsrude, 2003*; *Narain et al., 2003*; *Rodd et al., 2005*, *2010*; *Okada et al., 2010*; *Peelle et al., 2010*) by showing that temporal brain regions entrain to the natural rhythm of speech signal in both blind and sighted groups (*Figure 1F*). Strikingly, we show that following early visual deprivation, occipito-parietal regions enhance their synchronization to the acoustic rhythm of speech (*Figure 1G*) independently of language content. That is, the region around the Precuneus/Cuneus area seems to be involved in low-level processing of acoustic informa-tion alone. Previous studies have demonstrated that this area enhances its response to auditory information after early, but not late, visual deprivation (*Collignon et al., 2013*) and maintains this elevated response to sound even after sight-restoration early in life (*Collignon et al., 2015*). These studies did not find that this crossmodal response was related to a specific cognitive function, but rather pointed to a more general response to sounds (*Collignon et al., 2013*, *2015*). Interestingly, the current data suggest that this area is capable of parsing complex temporal patterns in sounds
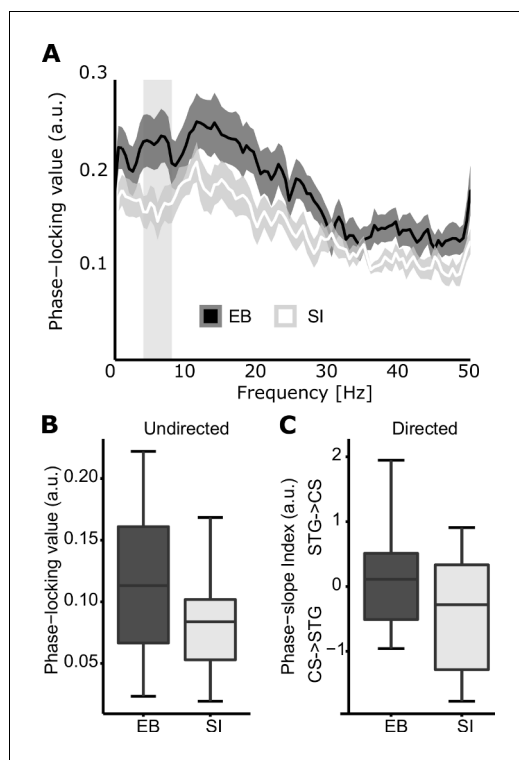
Figure 2. Occipital-temporal connectivity in EB and SI. (A) Spectra illustrate the phase locking between CS and STG in EB (dark curve) and in SI (light curve). Shaded areas illustrate the SEM. A difference between the curves is observed in the theta range. (B) Boxplots depict the mean phase locking between CS and STG for EB (black) and SI (white) in the theta range. Connectivity is enhanced in EB versus SI. (C) Directional connectivity analysis using the phase-slope index (PSI). Positive values suggest enhanced directionality from STG to CS, and negative values represent enhanced connectivity in the opposite direction. The boxplots highlight the strong feed-forward drive from CS to STG shown by SI, whereas blind individuals show a more balanced pattern, with a non-significant trend in the opposite direction.

DOI: https://doi.org/10.7554/eLife.31640.005

but is insensitive to the intelligibility of speech, again suggesting a more general role in sound processing.

In order to isolate the brain regions that are modulated by speech comprehension, we identified regions showing a greater response to amplitude modulations that convey speech information as compared to amplitude modulations that do not. In addition to the right STG observed in both groups (*Figure 1H*), the blind showed enhanced cerebro-acoustic coherence during intelligible speech in the vicinity of calcarine sulcus (V1; *Figure 1I*). This pattern of local encoding was accompanied by enhanced occipito-temporal connectivity during speech comprehension in EB as compared to SI. SI show the expected feed-forward projections from occipital to temporal regions (*Lamme et al., 1998*), whereas EB show a more balanced connectivity profile, trending towards the reverse temporal to occipital direction (see *Figure 2*). These findings support the idea of a reverse hierarchical model (*Büchel, 2003*) of the occipital cortex in EB, where the regions typically coding for 'low-level' visual features in the sighted (e.g. visual contrast or orientation) participate in higher-level function (e.g. speech intelligibility). Indeed, previous studies have found increased activity in the primary 'visual' cortex of EB people during Braille reading (*Sadato et al., 1996*; *Burton et al., 2002*, *2012*), verbal memory and verb generation tasks (*Amedi et al., 2003*), and during auditory language-related processing (*Bedny et al., 2011*). In line with our results, activity in primary occipital regions in EB people is stronger in a semantic versus a phonologic task (*Burton et al., 2003*), and vary as a function of syntactic and semantic complexity (*Röder et al., 2002*; *Bedny et al., 2011*; *Lane et al., 2015*). Moreover, repetitive transcranial magnetic stimulation (rTMS) over the occipital pole induces more semantic errors than phonologic errors in a verb-generation task in EB people (*Amedi et al.,*

2003). As we show that occipital regions entrain to the envelope of speech and are enhanced by its intelligibility, our results clearly suggest that the involvement of the occipital pole for language is not fully abstracted from sensory inputs as previously suggested (*Bedny, 2017*). The role of the occipital cortex in tracking the flow of speech in EB people may constitute an adaptive strategy to boost perceptual sensitivity at informational peaks in language.

Mechanistic origins of envelope tracking in occipital cortex

Although language is not the only cognitive process that selectively activates the occipital cortex of EB people, it is arguably one of the most puzzling. The reason is that reorganization of other processes, such as auditory motion perception and tactile object recognition, appears to follow the topography of the functionally equivalent visual processes in the sighted brain (*Ricciardi et al., 2007a*; *Amedi et al., 2010*; *Dormal et al., 2016*). For example, the hMT+/V5 complex, which is typically involved in visual motion in the sighted, selectively processes auditory

(*Poirier et al., 2006*; *Dormal et al., 2016*; *Jiang et al., 2016*) or tactile (*Ricciardi et al., 2007b*) motion in blind people. However, in the case of language, such recruitment is striking in light of the cognitive and evolutionary differences between vision and language (*Bedny et al., 2011*). This led to the proposal that, at birth, human cortical areas are cognitively pluripotent: capable of assuming a broad range of unrelated cognitive functions (*Bedny, 2017*). However, this argument resides on the presupposition that language has no computational relation with vision. But does this proposition tally with what we know about the relationship between the visual system and the classical language network? Rhythmic information in speech has a well-known language-related surrogate in the visual domain: lip movements (*Lewkowicz and Hansen-Tift, 2012*; *Park et al., 2016*). Indeed, both acoustic and visual speech signals exhibit rhythmic temporal patterns at prosodic and syllabic rates (*Chandrasekaran et al., 2009*; *Schwartz and Savariaux, 2014*; *Giordano et al., 2017*). The perception of lip kinematics that are naturally linked to amplitude fluctuations in speech serves as an important vehicle for the everyday use of language (*Kuhl and Meltzoff, 1982*; *Weikum et al., 2007*; *Lewkowicz and Hansen-Tift, 2012*) and helps language understanding, particularly in noisy conditions (*Ross et al., 2007*). Indeed, reading lips in the absence of any sound activates both primary and association auditory regions overlapping with regions that are active during the actual perception of spoken words (*Calvert et al., 1997*). The synchronicity between auditory and visual speech entrains rhythmic activity in the observer's primary auditory and visual regions, and facilitates perception by aligning neural excitability with acoustic or visual speech features (*Schroeder et al., 2008*; *Schroeder and Lakatos, 2009*; *Giraud and Poeppel, 2012*; *Mesgarani and Chang, 2012*; *Peelle and Davis, 2012*; *van Wassenhove, 2013*; *Zion Golumbic et al., 2013b*; *Park et al., 2016*; *Giordano et al., 2017*). These results strongly suggest that both the auditory and the visual components of speech are processed together at the earliest level possible in neural circuitry, based on the shared slow temporal modulations (around 2–7 Hz range) present across modalities (*Chandrasekaran et al., 2009*). Corroborating this idea, it has been demonstrated that neuronal populations in visual cortex follow the temporal dynamics of lip movements in sighted individuals, similar to the way in which temporal regions follow the acoustic and visual fluctuations of speech (*Luo et al., 2010*; *Zion Golumbic et al., 2013a*; *Park et al., 2016*). Similar to temporal cortex, occipital cortex in the sighted also shows enhanced lip tracking when attention is directed to speech content. This result highlights the fact that a basic oscillatory architecture for tracking the dynamic aspects of (visual-) speech in occipital cortex exists even in sighted individuals.

Importantly, audiovisual integration of the temporal dynamics of speech has been suggested to play a key role when learning speech early in life: young infants detect, match, and integrate the auditory and visual temporal coherence of speech (*Kuhl and Meltzoff, 1982*; *Rosenblum et al., 1997*; *Lewkowicz, 2000*, *2010*; *Brookes et al., 2001*; *Patterson and Werker, 2003*; *Lewkowicz and Ghazanfar, 2006*; *Kushnerenko et al., 2008*; *Bristow et al., 2009*; *Pons et al., 2009*; *Vouloumanos et al., 2009*; *Lewkowicz et al., 2010*; *Nath et al., 2011*). For instance, young infants between 4 and 6 months of age can detect their native language from lip movements only (*Weikum et al., 2007*). Around the same period, children detect synchrony between lip movements and speech sounds, and distribute more attention towards the mouth than towards the eyes (*Lewkowicz and Hansen-Tift, 2012*). Linking what they hear to the lip movements may provide young infants with a stepping-stone towards language production (*Lewkowicz and Hansen-Tift, 2012*). Moreover, infants aged 10 weeks already exhibit a McGurk effect, again highlighting the early multisensory nature of speech perception (*Rosenblum et al., 1997*). Taken together, these results suggest that an audio-visual link between observing lip movements and hearing speech sounds is present at very early developmental stages, potentially from birth, which helps infants acquire their first language.

In line with these prior studies, the current results may support the biased connectivity hypothesis of cross-modal reorganization (*Reich et al., 2011*; *Hannagan et al., 2015*; *Striem-Amit et al., 2015*). Indeed, it has been argued that reorganization in blind occipital cortex may be constrained by functional pathways to other sensory and cognitive systems that are also present in sighted individuals (*Elman et al., 1996*; *Hannagan et al., 2015*). This hypothesis may explain the overlap in functional specialization between blind and sighted individuals (*Collignon et al., 2011*; *Dormal et al., 2016*; *He et al., 2013*; *Jiang et al., 2016*; *Peelen et al., 2013*; *Pietrini et al., 2004*; *Weeks et al., 2000*; *Poirier et al., 2006*). In our experiment, sensitivity to acoustic dynamics of intelligible speech in blind occipital cortex could arise from pre-existing occipito-temporal pathways connecting the

auditory and visual system that are particularly important for the early developmental stages of language acquisition. In fact, the reorganization of this potentially predisposed pathway to process language content would explain how language-selective response may appear in blind children as young as 3 years old (*Bedny et al., 2015*).

Previous studies have suggested that language processing in the occipital cortex arises through top-down projections from frontal regions typically associated with the classical language network (*Bedny et al., 2011*; *Deen et al., 2015*), and that the representational content is symbolic and abstract rather than sensory (*Bedny, 2017*). Our results contrast with this view by showing that neuronal populations in (peri-)calcarine cortex align to the temporal dynamics of intelligible speech, and are functionally connected to areas sensitive to auditory information in temporal cortex. In sighted individuals, regions of the temporal lobe including STS are sensitive to acoustic features of speech, whereas higher-level regions such as anterior temporal cortex and left inferior frontal gyrus are relatively insensitive to these features and therefore do not entrain to the syllabic rate of speech (*Davis and Johnsrude, 2003*; *Hickok and Poeppel, 2007*; see confirmation in *Figure 1F–G*). This suggests that occipital areas respond, at least partially, to speech at a much lower (sensory) level than previously thought in EB brains, which may be caused by the reorganization of existing multisensory pathways connecting the 'auditory' and 'visual' centres in the brain.

Functional dependencies between sensory systems exist between the earliest stages of sensory processing in both human (*Ghazanfar and Schroeder, 2006*; *Kayser et al., 2008*; *Schroeder and Lakatos, 2009*; *Murray et al., 2016*) and nonhuman primates (*Falchier et al., 2002*; *Lakatos et al., 2007*; *Schroeder and Lakatos, 2009*). Several neuroimaging studies have demonstrated enhanced functional connectivity in sighted individuals between auditory and visual cortices under multisensory conditions (see *Murray et al., 2016* for a recent review), including multisensory speech (*Giordano et al., 2017*). Moreover, neuroimaging studies have shown that hearing people consistently activate left temporal regions during silent speech-reading (*Calvert et al., 1997*; *MacSweeney et al., 2000*, *2001*). We therefore postulate that brain networks that are typically dedicated to the integration of audio-visual speech signal, might be reorganized in the absence of visual inputs and might lead to an extension of speech tracking in the occipital cortex (*Collignon et al., 2009*). Although an experience-dependent mechanism related to EB affects the strength and directionality of the connectivity between the occipital and temporal regions, the presence of intrinsic connectivity between these regions — which can also be observed in sighted individuals (e.g. for multisensory integration) — may constrain the expression of the plasticity observed in our task. Building on this connectivity bias, the occipital pole may extend its sensitivity to the intelligibility of speech, a computation this region is obviously not originally dedicated to.

A number of recent studies have suggested that visual deprivation reinforces the functional connections between the occipital cortex and auditory regions typically classified as the language network (*Hasson et al., 2016*; *Schepers et al., 2012*). Previous studies using dynamic causal modelling support the idea that auditory information reaches the reorganized occipital cortex of the blind through direct temporo-occipital connection, rather than using subcortical (*Klinge et al., 2010*) or top-down pathways (*Collignon et al., 2013*). In support of these studies, we observed that the overall magnitude of functional connectivity between occipital and temporal cortex is higher in blind people than in sighted people during natural speech comprehension. Moreover, directional connectivity analysis revealed that the interaction between the two cortices is also qualitatively different: sighted individuals show a strong feed-forward drive towards temporal cortex, whereas blind individuals show a more balanced information flow, and a trend in the reverse direction. These results highlight one possible pathway by which the speech signal is enhanced in the occipital cortex of EB individuals. However, this does not mean that the changes in connectivity between blind and sighted individuals are limited to this specific network (e.g. see *Kayser et al., 2015*; *Park et al., 2015*).

Speech comprehension in the right hemisphere

We observed that neuronal populations in right superior temporal cortex synchronize to the temporal dynamics of intelligible, but not non-intelligible, speech in both EB and SI groups. Why does speech intelligibility modulate temporal regions of the right, but not the left, hemisphere? According to an influential model of speech comprehension – the asymmetric sampling in time model (AST; *Giraud and Poeppel, 2012*; *Hickok and Poeppel, 2007*; *Poeppel, 2003*) – there is a division of

labour between the left- and right auditory cortices (*Poeppel, 2003*; *Boemio et al., 2005*; *Hickok and Poeppel, 2007*), with the left auditory cortex being more sensitive to high-frequency information (+20 Hz), whereas the right temporal cortex is more sensitive to low-frequency information (~6 Hz) such as syllable sampling and prosody (*Belin et al., 1998*; *Poeppel, 2003*; *Boemio et al., 2005*; *Obleser et al., 2008*; *Giraud and Poeppel, 2012*). Several studies have shown that the right hemisphere is specifically involved in the representation of connected speech (*Bourguignon et al., 2013*; *Fonteneau et al., 2015*; *Horowitz-Kraus et al., 2015*; *Alexandrou et al., 2017*), whereas other studies have directly demonstrated the prevalence of speech-to-brain entrainment while listening to sentences or stories in delta and theta bands in the right hemisphere more than in the left hemisphere (*Luo and Poeppel, 2007*; *Abrams et al., 2008*; *Gross et al., 2013*; *Giordano et al., 2017*). The present study therefore replicates these results by showing enhanced phase coupling between the right hemisphere and the speech envelope at the syllabic rate (low-frequency phase of speech envelope), consistent with the AST model. An interesting observation in the current study is that right hemispheric sensitivity to intelligible speech in temporal areas coincides with the enhanced right hemispheric sensitivity to intelligible speech in the occipital cortex of blind individuals.

Having more cortical tissue devoted to sentence processing and understanding could potentially support enhanced sentence comprehension. Previous studies have indeed demonstrated that, as compared to sighted individuals, blind people have enhanced speech discrimination in noisy environments (*Niemeyer and Starlinger, 1981*), as well as the capability to understand speech displayed at a much faster rate (sighted listeners at rates of 9–14 syllables/s andblind listeners at rates of 17–22 syllables/s; *Moos and Trouvain, 2007*). Crucially, listening to intelligible ultra-fast speech (as compared to reverse speech) has been shown to cause enhanced activity in the right primary visual cortex in early and late blind individuals when compared to sighted controls, and activity in this region correlates with individual ultra-fast speech perception skills [*Dietrich et al., 2013*]). These results raise the interesting possibility that the engagement of right V1 in the analysis of the speech envelope, as demonstrated in our study, may support the enhanced encoding of early supra-segmental aspects of the speech signal, supporting the ability to understand an ultra-fast speech signal. However, our behavioral task (speech comprehension) did not allow us to directly assess this link between the reorganized occipital cortex and speech comprehension, as by design performance was almost at ceiling in the nat and 8-chan condition but at chance in the 1-chan condition (see *Table 1*).

However, it is important to note that linking brain activity and behavior is not a trivial issue. Many studies have not found a direct link between crossmodal reorganization and non-visual processing. More generally, as a behavioral outcome is the end product of a complex interactive process between several brain regions, linking the role of one region in isolation (e.g. the reorganized region in the occipital cortex of blind people) to behavior (e.g. performance) in a multifaced task is not straightforward. An absence of a direct relation between behavior (e.g. speech processing) and crossmodal plasticity could be explained by the fact that this complex process is supported by additional networks other than the reorganized ones. Interestingly, recent studies have shown that early visual deprivation triggers a game of 'balance' between the brain systems typically dedicated to a specific process and the reorganized occipital cortex (*Dormal et al., 2016*). It has indeed been proposed that early visual deprivation triggers a redeployment mechanism that would reallocate part of the processing typically implemented in the preserved networks (i.e. the temporal or frontal cortices for speech processing) to the occipital cortex deprived of its most salient input (vision). Two recent studies using multivoxel pattern analysis (MVPA) showed that the ability to decode the different auditory motion stimuli was enhanced in hMT+ (a region typically involved in visual motion in sighted individuals) of early blind individuals, whereas an enhanced decoding accuracy was observed in the planum temporale in the sighted group (*Jiang et al., 2016*; *Dormal et al., 2016*). Moreover, *Bedny et al. (2015)* reported an enhanced activation of occipital cortex and a simultaneous deactivation of prefrontal regions during a linguistic task in blind children. The authors suggested that the increased involvement of the occipital cortex might decrease the pressure on the prefrontal areas to specialize in language processing. Interestingly, a transcranial magnetic stimulation (TMS) study reported a reduced disruptive effect of TMS applied over the left inferior prefrontal cortex during linguistic tasks for both blind and sighted individuals, whereas TMS applied over the occipital cortex caused more disruption in early blind as compared to sighted people (*Amedi et al., 2004*). Such

studies support the idea that brain regions that are typically recruited for specific tasks in sighted individuals may become less essential in EB people if they concomitantly recruit occipital regions for the same task. An important open question for future research therefore concerns the relative behavioral contribution of occipital and perisylvian cortex to speech understanding.

Conclusions

We demonstrate that the primary 'visual' cortex synchronizes to the temporal dynamics of intelligible speech at the rate of syllable transitions in language (~6–7 Hz). Our results demonstrate that this neural population is involved in processing the sensory signal of speech, and therefore contrasts with the proposition that occipital involvement in speech processing is abstracted from its sensory input and purely reflects higher-level operations similar to those observed in prefrontal regions (*Bedny, 2017*). Blindness, due to the absence of organizing visual input, leaves the door open for sensory and functional colonization of occipital regions. This colonization might however not be stochastic, but could be constrained by modes of information processing that are natively anchored in specific brain regions and networks. Even if the exact processing mode is still to be unveiled by future research, we hypothesise that the mapping of language onto occipital cortex builds upon pre-existing oscillatory architecture typically linking auditory and visual speech rhythm (*Chandrasekaran et al., 2009*). Our study therefore supports the view that the development of functional specialisation in the human cortex is the product of a dynamic interplay between genetic and environmental factors during development, rather than being predetermined at birth (*Elman et al., 1996*).

Materials and methods

Participants

Seventeen early blind (11 female; mean ± SD, 32.9 ± 10.19 years; range, 20–67 years) and sixteen sighted individuals (10 female; mean ± SD, 32.2 ± 9.92 years; range, 20–53 years) participated in the current study. There was no age difference between the blind and the sighted group (t[30] = 0.19, p=0.85). All blind participants were either totally blind or severely visually impaired from birth; however, two participants reported residual visual perception before the age of 3 and one before the age of 4, as well as one participant who lost their sight completely at age 10. Causes of vision loss were damage or detached retina (10), damage to the optic nerve (3), infection of the eyes (1), microphtalmia (2), and hypoxia (1). Although some participants reported residual light perception, none were able to use vision functionally. All participants were proficient braille readers and native speakers of Italian. None of them suffered from a known neurological or peripheral auditory disease. The data from one sighted individual were not included because of the discovery of a brain structural abnormality that was unknown to the experimenters at the time of the experiment. The project was approved by the local ethical committee at the University of Trento. In agreement with the Declaration of Helsinki, all participants provided written informed consent before participating in the study.

Experimental design

Auditory stimuli were delivered into the magnetically shielded MEG room via stereo loudspeakers using a Panphonics Sound Shower two amplifier at a comfortable sound level, which was the same for all participants. Stimulus presentation was controlled via the Matlab (RRID:SCR_001622) Psychophysics Toolbox 3 (http://psychtoolbox.org; RRID:SCR_002881) running on a Dell Alienware Aurora PC under Windows 7 (64 bit). Both sighted and blind participants were blindfolded during the experiment, and the room was dimly lit to allow for visual monitoring of the participant via a video feed from the MEG room. Instructions were provided throughout the experiment using previous recordings from one of the experimenters (FB).

The stimuli consisted of 14 short segments (~1 min) from popular audiobooks (e.g., Pippi Longstocking and Candide) in Italian (nat). Furthermore, channel-vocoding in the Praat software was used to produce two additional control conditions. First, the original sound file was band-pass filtered into 1 (1-channel) or 8 (8-channel) logarithmically spaced frequency bands. The envelope for each of these bands was computed, filled with Gaussian white noise, and the different signals were recombined into a single sound file. The resulting signal has an amplitude envelope close to the original,

while the fine spectral detail was gradually distorted. Perceptually, the voice of the speaker is highly distorted in the 8-channel condition, but intelligibility is unperturbed. By contrast, in the 1-channel condition, speech is entirely unintelligible. In total, 42 sound files were presented in a pseudo-randomized fashion, distributed among seven blocks. Each block contained two sound files from each condition, and the same story was never used twice in the same block. Examples of the stimuli are provided in supplementary media content (see *Video 1*).

To verify story comprehension, each speech segment was followed by a single-sentence statement about the story. Participants were instructed to listen to the story carefully, and to judge whether the statement at the end was true or false, using a nonmagnetic button box. Responses were provided with the index and middle finger of the right hand.

MEG data acquisition and preprocessing

MEG was recorded continuously from a 306 triple sensor (204 planar gradiometers; 102 magnetometers) whole-head system (Elekta Neuromag, Helsinki, Finland) using a sampling rate of 1 kHz and an online band-bass filter between 0.1 and 300 Hz. The headshape of each individual participant was measured using a Polhemus FASTRAK 3D digitizer. Head position of the subject was recorded continuously using five localization coils (forehead, mastoids).

Data pre-processing was performed using the open-source Matlab toolbox Fieldtrip (www.fieldtriptoolbox.org; RRID:SCR_004849). First the continuous data were filtered (high-pass Butterworth filter at 1 Hz, low-pass Butterworth filter at 170 Hz, and DFT filter at 50,100, and 150 Hz to remove line-noise artefacts in the signal), and downsampled to 256 Hz. Next, the data were epoched into segments of 1 s for subsequent analysis.

Rejection of trials containing artefacts and bad channels was performed using a semi-automatic procedure. First, outliers were rejected using a pre-screening based on the variance and range in each trial/channel. Then, algorithmically guided visual inspection of the raw data was performed to remove any remaining sources of noise.

Extraction of the speech amplitude envelope

Amplitude envelopes of the stories were computed using the Chimera toolbox (*Chandrasekaran et al., 2009*) and custom code, following the procedure described by Gross and colleagues (*Gross et al., 2013*). First, the sound files were band-pass filtered between 100 and 1000 Hz into nine frequency-bands, using a fourth order Butterworth filter. The filter was applied in forward and backward direction to avoid any phase shifts with respect to the original signal, and frequency-bands were spaced with equal width along the human basilar membrane. Subsequently, the analytic amplitude for each filtered segment was computed as the absolute of the Hilbert transform. Finally, the amplitude envelopes of all nine bands were summed and scaled to a maximum value of 1. The resulting envelope was combined with the MEG data, and processed identically henceforth.

Analysis of cerebro-acoustic coherence

To determine where, and at what temporal scale, neuronal populations follow the temporal dynamics of speech, we computed spectral coherence between the speech envelope and the MEG signal. Coherence is a statistic that quantifies the phase relationship between two signals, and can be used to relate oscillatory activity in the brain with a peripheral measure such as a speech signal (*Peelle et al., 2013*). The first analysis was conducted in sensor space, across conditions and participants, to determine the temporal scale at which coherence between the speech envelope and the MEG signal is strongest. To this end, a Hanning taper was applied to the 1 s data segments, and Fourier transformation was used to compute the cross-spectral density between 1 and 30 Hz, with a step size of 1 Hz. To render the coherence values more normally distributed, a Fisher z-transform was applied by computing the inverse hyperbolic tangent (atanh).

Source-space analysis was centred on the coherence frequency-band of interest (FOI) identified in the sensor space analysis. Source reconstruction was performed using a frequency domain beamformer called Dynamic Imaging of Coherent Sources (DICS) (*Gross et al., 2001*; *Liljeström et al., 2005*). DICS was used to localize spectral coherence, observed at the sensors, on a three-dimensional grid (8 × 8 × 8 mm). The forward model was based on a realistic single-shell headmodel (*Nolte, 2003*) for each individual. As structural scans could not be acquired for all participants, we

approximated individual anatomy by warping a structural MNI template brain (MNI, Montreal, Quebec, Canada; www.bic.mni.mcgill.ca/brainweb) into individual headspace using the information from each participant's headshape.

Source connectivity analysis

Functional connectivity between occipital and temporal cortex was computed by extracting virtual sensor time-series at the locations of interest in CS and STG using time-domain beamforming. These virtual sensor time series were used to compute non-directional and directional connectivity metrics. The rationale behind focusing our analyses on STG and CS is to limit our hypothesis space by the results as they unfolded in our analytic steps. Indeed, we found that the right STG was the only region modulated by the intelligibility of speech in both groups, confirming previous results using similar methods (*Luo and Poeppel, 2007*; *Abrams et al., 2008*; *Gross et al., 2013*; *Giordano et al., 2017*). In addition to STG, we found enhanced cerebro-acoustic coherence for intelligible speech in CS of EB individuals. We therefore decided to focus our connectivity analyses on these two ROIs since they were functionally defined from our first analytical step (modulation of cerebro-acoustic coherence by speech intelligibility). The advantage of following this procedure is that connectivity analyses are based on nodes that we can functionally interpret and for which we have clear predictions. We successfully used the same hierarchical analytic structure in previous studies (*Collignon et al., 2013*, *2015*; *Benetti et al., 2017*), following guidelines on how to investigate directional connectivity in brain networks (*Stephan et al., 2010*).

Virtual sensor time-series at the two locations of interest were extracted using a time-domain vector-beamforming technique called linear constrained minimum variance (LCMV) beamforming (*Van Veen et al., 1997*). First, average covariance matrices were computed for each participant to estimate common spatial filter coefficients. These filter coefficients were multiplied with the single-trial cleaned MEG data. To reduce the resulting three-dimensional time-series to one singular value, decomposition (SVD) was applied, resulting in a single time-series for each trial and participant.

Functional connectivity between the two regions was computed at peak locations using the phase-locking value (PLV) (*Lachaux et al., 1999*). First, single-trial virtual sensor time courses were converted to the frequency domain (0–50 Hz) using Fourier transformation. The data were padded to 2 s and a Hanning taper was applied to reduce spectral leakage. Coherence was used as a proxy for functional connectivity. To disentangle phase consistency between regions from joint fluctuations in power, the spectra were normalized with respect to the amplitude, resulting in an estimate of the phase-locking (*Lachaux et al., 1999*) between regions. Connectivity estimates were normalized using a Fischer z-transform (atanh), as in the analysis of cerebro-acoustic coherence.

PLV is a symmetric proxy for connectivity and does not allow for inferences regarding the direction of information flow between two regions of interest. To test whether differences in functional connectivity between EB and SI are accompanied by changes in the directionality of the flow of information between the regions of interest, we computed the phase slope index (PSI) (*Nolte et al., 2008*). The PSI deduces net directionality from the time delay between two time series (x1 and x2). Time series x1 is said to precede, and hence drive, time series x2 in a given frequency band if the phase difference between x1 and x2 increases with higher frequencies. Consequently, a negative phase slope reflects a net information flux in the reverse direction, that is, from x2 to x1. Here, we computed the PSI using a bandwidth of ± 5 Hz around the frequencies of interest. Following the recommendation by Nolte and colleagues (*Nolte et al., 2008*), PSI estimates were normalized with the standard error, which was computed using the jackknife method.

Statistical analysis

Statistical testing of the behavioural comprehension scores, as well as the connectivity estimates, was performed using linear mixed-effects models in R. Differences between conditions and groups in source space were evaluated using Statistical Parametric Mapping (SPM12; RRID:SCR_007037), and the Threshold-Free Cluster Enhancement (TFCE) toolboxes in Matlab. TFCE (*Smith and Nichols, 2009*) computes new values for each voxel in a statistical map as a function of the original voxel value and the values of the surrounding voxels. By enhancing the statistical values in voxels with a high T-value that are also part of a local cluster, TFCE optimally combines the benefits of voxel-wise and cluster-based methods. TFCE was applied to the whole-brain contrasts. Final correction for

multiple comparisons was applied using a maximum statistic based on 1000 permutations of group membership (independent testing) or conditions (dependent testing). Here, we applied a variance smoothing of 15 mm FWHM to reduce the effects of high spatial frequency noise in the statistical maps. The smoothing kernel used in the current study is higher than that in comparable fMRI studies because of the inherent smoothness of the source-reconstructed MEG data.

Acknowledgements

Additional information

Author contributions

Markus Johannes Van Ackeren, Conceptualization, Data curation, Software, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Francesca M Barbero, Conceptualization, Data curation, Validation, Methodology, Project administration, Writing—review and editing; Stefania Mattioni, Roberto Bottini, Resources, Data curation, Validation, Investigation, Project administration, Writing—review and editing; Olivier Collignon, Conceptualization, Resources, Data curation, Supervision, Funding acquisition, Validation, Visualization, Methodology, Writing—original draft, Project administration, Writing—review and editing

Author ORCIDs

Francesca M Barbero (iD) https://orcid.org/0000-0003-4445-4402
Roberto Bottini (iD) http://orcid.org/0000-0001-7941-7762
Olivier Collignon (iD) http://orcid.org/0000-0003-1882-3550

Ethics

Human subjects: The project was approved by the local ethical committee at the University of Trento (protocol 2014-007). In agreement with the Declaration of Helsinki, all participants provided written informed consent to participate in the study.

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.31640.009
Author response https://doi.org/10.7554/eLife.31640.010

Additional files

References

Abrams DA, Nicol T, Zecker S, Kraus N. 2008. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *Journal of Neuroscience* **28**:3958–3965. DOI: https://doi.org/10.1523/JNEUROSCI.0187-08.2008, PMID: 18400895

Alexandrou AM, Saarinen T, Mäkelä S, Kujala J, Salmelin R. 2017. The right hemisphere is highlighted in connected natural speech production and perception. *NeuroImage* **152**:628–638. DOI: https://doi.org/10.1016/j.neuroimage.2017.03.006, PMID: 28268122

Amedi A, Floel A, Knecht S, Zohary E, Cohen LG. 2004. Transcranial magnetic stimulation of the occipital pole interferes with verbal processing in blind subjects. *Nature Neuroscience* **7**:1266–1270. DOI: https://doi.org/10.1038/nn1328, PMID: 15467719

Amedi A, Raz N, Azulay H, Malach R, Zohary E. 2010. Cortical activity during tactile exploration of objects in blind and sighted humans. *Restorative Neurology and Neuroscience* **28**:143–156. DOI: https://doi.org/10.3233/RNN-2010-0503, PMID: 20404404

Amedi A, Raz N, Pianka P, Malach R, Zohary E. 2003. Early 'visual' cortex activation correlates with superior verbal memory performance in the blind. *Nature Neuroscience* **6**:758–766. DOI: https://doi.org/10.1038/nn1072, PMID: 12808458

Arnaud L, Sato M, Ménard L, Gracco VL. 2013. Repetition suppression for speech processing in the associative occipital and parietal cortex of congenitally blind adults. *PLoS One* **8**:e64553. DOI: https://doi.org/10.1371/journal.pone.0064553, PMID: 23717628

Atilgan H, Collignon O, Hasson U. 2017. Structural neuroplasticity of the superior temporal plane in early and late blindness. *Brain and Language* **170**:71–81. DOI: https://doi.org/10.1016/j.bandl.2017.03.008, PMID: 28426947

Bates E. 1999. Language and the infant brain. *Journal of Communication Disorders* **32**:195–205. DOI: https://doi.org/10.1016/S0021-9924(99)00015-5, PMID: 10466093

Bates ED. 2005. Plasticity, Localization, and Language Development. In: Milbrath C. onstance (Ed). *Biol Knowl Revisit From Neurogenes to Psychogenes*. **9351** p. 205–253.

Bedny M, Pascual-Leone A, Dodell-Feder D, Fedorenko E, Saxe R. 2011. Language processing in the occipital cortex of congenitally blind adults. *PNAS* **108**:4429–4434. DOI: https://doi.org/10.1073/pnas.1014818108, PMID: 21368161

Bedny M, Richardson H, Saxe R. 2015. "Visual" cortex responds to spoken language in blind children. *The Journal of Neuroscience* **35**:11674–11681. DOI: https://doi.org/10.1523/JNEUROSCI.0634-15.2015, PMID: 26290244

Bedny M. 2017. Evidence from blindness for a cognitively pluripotent cortex. *Trends in Cognitive Sciences* **21**:637–648. DOI: https://doi.org/10.1016/j.tics.2017.06.003, PMID: 28821345

Belin P, Zilbovicius M, Crozier S, Thivard L, Fontaine A, Masure MC, Samson Y. 1998. Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience* **10**:536–540. DOI: https://doi.org/10.1162/089892998562834, PMID: 9712682

Benetti S, van Ackeren MJ, Rabini G, Zonca J, Foa V, Baruffaldi F, Rezk M, Pavani F, Rossion B, Collignon O. 2017. Functional selectivity for face processing in the temporal voice area of early deaf individuals. *PNAS* **114**:E6437–E6446. DOI: https://doi.org/10.1073/pnas.1618287114, PMID: 28652333

Berwick RC, Friederici AD, Chomsky N, Bolhuis JJ. 2013. Evolution, brain, and the nature of language. *Trends in Cognitive Sciences* **17**:89–98. DOI: https://doi.org/10.1016/j.tics.2012.12.002, PMID: 23313359

Boemio A, Fromm S, Braun A, Poeppel D. 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience* **8**:389–395. DOI: https://doi.org/10.1038/nn1409, PMID: 15723061

Bourguignon M, De Tiège X, de Beeck MO, Ligot N, Paquier P, Van Bogaert P, Goldman S, Hari R, Jousmäki V. 2013. The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Human Brain Mapping* **34**:314–326. DOI: https://doi.org/10.1002/hbm.21442, PMID: 22392861

Bristow D, Dehaene-Lambertz G, Mattout J, Soares C, Gliga T, Baillet S, Mangin JF. 2009. Hearing faces: how the infant brain matches the face it sees with the speech it hears. *Journal of Cognitive Neuroscience* **21**:905–921. DOI: https://doi.org/10.1162/jocn.2009.21076, PMID: 18702595

Brookes H, Slater A, Quinn PC, Lewkowicz DJ, Hayes R, Brown E. 2001. Three-month-old infants learn arbitrary auditory-visual pairings between voices and faces. *Infant and Child Development* **10**:75–82. DOI: https://doi.org/10.1002/icd.249

Burton H, Diamond JB, McDermott KB. 2003. Dissociating cortical regions activated by semantic and phonological tasks: a FMRI study in blind and sighted people. *Journal of Neurophysiology* **90**:1965–1982. DOI: https://doi.org/10.1152/jn.00279.2003, PMID: 12789013

Burton H, Sinclair RJ, Agato A. 2012. Recognition memory for Braille or spoken words: an fMRI study in early blind. *Brain Research* **1438**:22–34. DOI: https://doi.org/10.1016/j.brainres.2011.12.032, PMID: 22251836

Burton H, Snyder AZ, Diamond JB, Raichle ME. 2002. Adaptive changes in early and late blind: a FMRI study of verb generation to heard nouns. *Journal of Neurophysiology* **88**:3359–3371. DOI: https://doi.org/10.1152/jn.00129.2002, PMID: 12466452

Burton H. 2003. Visual cortex activity in early and late blind people. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* **23**:4005–4011. PMID: 12764085

Büchel C. 2003. Cortical hierarchy turned on its head. *Nature Neuroscience* **6**:657–658. DOI: https://doi.org/10.1038/nn0703-657, PMID: 12830152

Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PW, Iversen SD, David AS. 1997. Activation of auditory cortex during silent lipreading. *Science* **276**:593–596. DOI: https://doi.org/10.1126/science.276.5312.593, PMID: 9110978

Chandrasekaran C, Trubanova A, Stillittano S, Caplier A, Ghazanfar AA. 2009. The natural statistics of audiovisual speech. *PLoS Computational Biology* **5**:e1000436. DOI: https://doi.org/10.1371/journal.pcbi.1000436, PMID: 19609344

Chomsky N. 1976. *Reflections on Language.*

Collignon O, Dormal G, Albouy G, Vandewalle G, Voss P, Phillips C, Lepore F. 2013. Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. *Brain* **136**:2769–2783. DOI: https://doi.org/10.1093/brain/awt176, PMID: 23831614

Collignon O, Dormal G, de Heering A, Lepore F, Lewis TL, Maurer D. 2015. Long-lasting crossmodal cortical reorganization triggered by brief postnatal visual deprivation. *Current Biology* **25**:2379–2383. DOI: https://doi.org/10.1016/j.cub.2015.07.036, PMID: 26299512

Collignon O, Vandewalle G, Voss P, Albouy G, Charbonneau G, Lassonde M, Lepore F. 2011. Functional specialization for auditory-spatial processing in the occipital cortex of congenitally blind humans. *PNAS* **108**:4435–4440. DOI: https://doi.org/10.1073/pnas.1013928108, PMID: 21368198

Collignon O, Voss P, Lassonde M, Lepore F. 2009. Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. *Experimental Brain Research* **192**:343–358. DOI: https://doi.org/10.1007/s00221-008-1553-z, PMID: 18762928

Davis MH, Johnsrude IS. 2003. Hierarchical processing in spoken language comprehension. *Journal of Neuroscience* **23**:3423–3431. PMID: 12716950

de Heer WA, Huth AG, Griffiths TL, Gallant JL, Theunissen FE. 2017. The hierarchical cortical organization of human speech processing. *The Journal of Neuroscience* **37**:6539–6557. DOI: https://doi.org/10.1523/JNEUROSCI.3267-16.2017, PMID: 28588065

Deen B, Saxe R, Bedny M. 2015. Occipital cortex of blind individuals is functionally coupled with executive control areas of frontal cortex. *Journal of Cognitive Neuroscience* **27**:1633–1647. DOI: https://doi.org/10.1162/jocn_a_00807, PMID: 25803598

Dehaene-Lambertz G, Hertz-Pannier L, Dubois J. 2006. Nature and nurture in language acquisition: anatomical and functional brain-imaging studies in infants. *Trends in Neurosciences* **29**:367–373. DOI: https://doi.org/10.1016/j.tins.2006.05.011, PMID: 16815562

Dietrich S, Hertrich I, Ackermann H. 2013. Ultra-fast speech comprehension in blind subjects engages primary visual cortex, fusiform gyrus, and pulvinar - a functional magnetic resonance imaging (fMRI) study. *BMC Neuroscience* **14**:74. DOI: https://doi.org/10.1186/1471-2202-14-74, PMID: 23879896

Ding N, Simon JZ. 2014. Cortical entrainment to continuous speech: functional roles and interpretations. *Frontiers in Human Neuroscience* **8**:311. DOI: https://doi.org/10.3389/fnhum.2014.00311, PMID: 24904354

Dormal G, Rezk M, Yakobov E, Lepore F, Collignon O. 2016. Auditory motion in the sighted and blind: Early visual deprivation triggers a large-scale imbalance between auditory and "visual" brain regions. *NeuroImage* **134**:630–644. DOI: https://doi.org/10.1016/j.neuroimage.2016.04.027, PMID: 27107468

Elman JL, Bates EA, Johnson MH, Karmiloff-Smith A, Parisi D, Plunkett K. 1996. *Rethinking Innateness: Connectionist Perspective on Development.*

Falchier A, Clavagnier S, Barone P, Kennedy H. 2002. Anatomical evidence of multimodal integration in primate striate cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* **22**:5749–5759. PMID: 12097528

Fonteneau E, Bozic M, Marslen-Wilson WD. 2015. Brain network connectivity during language comprehension: interacting linguistic and perceptual subsystems. *Cerebral Cortex* **25**:3962–3976. DOI: https://doi.org/10.1093/cercor/bhu283, PMID: 25452574

Ghazanfar AA, Schroeder CE. 2006. Is neocortex essentially multisensory? *Trends in Cognitive Sciences* **10**:278–285. DOI: https://doi.org/10.1016/j.tics.2006.04.008, PMID: 16713325

Giordano BL, Ince RAA, Gross J, Schyns PG, Panzeri S, Kayser C. 2017. Contributions of local speech encoding and functional connectivity to audio-visual speech perception. *eLife* **6**:e24763. DOI: https://doi.org/10.7554/eLife.24763, PMID: 28590903

Giraud AL, Poeppel D. 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience* **15**:511–517. DOI: https://doi.org/10.1038/nn.3063, PMID: 22426255

Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S. 2013. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology* **11**:e1001752. DOI: https://doi.org/10.1371/journal.pbio.1001752, PMID: 24391472

Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R, Hämäläinen M. 2001. Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *PNAS* **98**:694–699. DOI: https://doi.org/10.1073/pnas.98.2.694, PMID: 11209067

Hannagan T, Amedi A, Cohen L, Dehaene-Lambertz G, Dehaene S. 2015. Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends in Cognitive Sciences* **19**:374–382. DOI: https://doi.org/10.1016/j.tics.2015.05.006, PMID: 26072689

Hasson U, Andric M, Atilgan H, Collignon O. 2016. Congenital blindness is associated with large-scale reorganization of anatomical networks. *NeuroImage* **128**:362–372. DOI: https://doi.org/10.1016/j.neuroimage.2015.12.048, PMID: 26767944

He C, Peelen MV, Han Z, Lin N, Caramazza A, Bi Y. 2013. Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. *NeuroImage* **79**:1–9. DOI: https://doi.org/10.1016/j.neuroimage.2013.04.051, PMID: 23624496

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nature Reviews Neuroscience* **8**:393–402. DOI: https://doi.org/10.1038/nrn2113, PMID: 17431404

Horowitz-Kraus T, Grainger M, DiFrancesco M, Vannest J, Holland SK, CMIND Authorship Consortium. 2015. Right is not always wrong: DTI and fMRI evidence for the reliance of reading comprehension on language-comprehension networks in the right hemisphere. *Brain Imaging and Behavior* **9**:19–31. DOI: https://doi.org/10.1007/s11682-014-9341-9, PMID: 25515348

Jiang F, Stecker GC, Boynton GM, Fine I. 2016. Early blindness results in developmental plasticity for auditory motion processing within auditory and occipital cortex. *Frontiers in Human Neuroscience* **10**:324. DOI: https://doi.org/10.3389/fnhum.2016.00324, PMID: 27458357

Kayser C, Montemurro MA, Logothetis NK, Panzeri S. 2009. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* **61**:597–608. DOI: https://doi.org/10.1016/j.neuron.2009.01.008, PMID: 19249279

Kayser C, Petkov CI, Logothetis NK. 2008. Visual modulation of neurons in auditory cortex. *Cerebral Cortex* **18**:1560–1574. DOI: https://doi.org/10.1093/cercor/bhm187, PMID: 18180245

Kayser SJ, Ince RA, Gross J, Kayser C. 2015. Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *Journal of Neuroscience* **35**:14691–14701. DOI: https://doi.org/10.1523/JNEUROSCI.2243-15.2015, PMID: 26538641

Klinge C, Eippert F, Röder B, Büchel C. 2010. Corticocortical connections mediate primary visual cortex responses to auditory stimulation in the blind. *Journal of Neuroscience* **30**:12798–12805. DOI: https://doi.org/10.1523/JNEUROSCI.2384-10.2010, PMID: 20861384

Kuhl PK, Meltzoff AN. 1982. The bimodal perception of speech in infancy. *Science.* **218**:1138–1141. DOI: https://doi.org/10.1126/science.7146899, PMID: 7146899

Kushnerenko E, Teinonen T, Volein A, Csibra G. 2008. Electrophysiological evidence of illusory audiovisual speech percept in human infants. *PNAS* **105**:11442–11445. DOI: https://doi.org/10.1073/pnas.0804275105, PMID: 18682564

Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. 1999. Measuring phase synchrony in brain signals. *Human Brain Mapping* **8**:194–208. DOI: https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C, PMID: 10619414

Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. 2007. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* **53**:279–292. DOI: https://doi.org/10.1016/j.neuron.2006.12.011, PMID: 17224408

Lamme VA, Supèr H, Spekreijse H, Feedforward SH. 1998. Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology* **8**:529–535. DOI: https://doi.org/10.1016/S0959-4388(98)80042-1, PMID: 9751656

Lane C, Kanjlia S, Omaki A, Bedny M. 2015. "Visual" cortex of congenitally blind adults responds to syntactic movement. *Journal of Neuroscience* **35**:12859–12868. DOI: https://doi.org/10.1523/JNEUROSCI.1256-15.2015, PMID: 26377472

Lewkowicz DJ, Ghazanfar AA. 2006. The decline of cross-species intersensory perception in human infants. *PNAS* **103**:6771–6774. DOI: https://doi.org/10.1073/pnas.0602027103, PMID: 16618919

Lewkowicz DJ, Hansen-Tift AM. 2012. Infants deploy selective attention to the mouth of a talking face when learning speech. *PNAS* **109**:1431–1436. DOI: https://doi.org/10.1073/pnas.1114783109, PMID: 22307596

Lewkowicz DJ, Leo I, Simion F. 2010. Intersensory perception at birth: newborns match nonhuman primate faces and voices. *Infancy* **15**:46–60. DOI: https://doi.org/10.1111/j.1532-7078.2009.00005.x

Lewkowicz DJ. 2000. Infants' perception of the audible, visible, and bimodal attributes of multimodal syllables. *Child Development* **71**:1241–1257. DOI: https://doi.org/10.1111/1467-8624.00226, PMID: 11108094

Lewkowicz DJ. 2010. Infant perception of audio-visual speech synchrony. *Developmental Psychology* **46**:66–77. DOI: https://doi.org/10.1037/a0015579, PMID: 20053007

Liljeström M, Kujala J, Jensen O, Salmelin R. 2005. Neuromagnetic localization of rhythmic activity in the human brain: a comparison of three methods. *NeuroImage* **25**:734–745. DOI: https://doi.org/10.1016/j.neuroimage.2004.11.034, PMID: 15808975

Luo H, Liu Z, Poeppel D. 2010. Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biology* **8**:e1000445. DOI: https://doi.org/10.1371/journal.pbio.1000445, PMID: 20711473

Luo H, Poeppel D. 2007. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* **54**:1001–1010. DOI: https://doi.org/10.1016/j.neuron.2007.06.004, PMID: 17582338

MacSweeney M, Amaro E, Calvert GA, Campbell R, David AS, McGuire P, Williams SC, Woll B, Brammer MJ. 2000. Silent speechreading in the absence of scanner noise: an event-related fMRI study. *Neuroreport* **11**:1729–1733 . DOI: https://doi.org/10.1097/00001756-200006050-00026, PMID: 10852233

MacSweeney M, Campbell R, Calvert GA, McGuire PK, David AS, Suckling J, Andrew C, Woll B, Brammer MJ. 2001. Dispersed activation in the left temporal cortex for speech-reading in congenitally deaf people. *Proceedings of the Royal Society B: Biological Sciences* **268**:451–457. DOI: https://doi.org/10.1098/rspb.2000.0393, PMID: 11296856

Mesgarani N, Chang EF. 2012. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* **485**:233–236. DOI: https://doi.org/10.1038/nature11020, PMID: 22522927

Moos A, Trouvain J. 2007. Comprehension of ultra-fast speech–blind vs.'normally hearing' persons. *Proc 16th Int Congr Phonetic Sci* 677–680.

Murray MM, Lewkowicz DJ, Amedi A, Wallace MT. 2016. Multisensory processes: a balancing act across the lifespan. *Trends in Neurosciences* **39**:567–579. DOI: https://doi.org/10.1016/j.tins.2016.05.003, PMID: 27282408

Narain C, Scott SK, Wise RJ, Rosen S, Leff A, Iversen SD, Matthews PM. 2003. Defining a left-lateralized response specific to intelligible speech using fMRI. *Cerebral Cortex* **13**:1362–1368. DOI: https://doi.org/10.1093/cercor/bhg083, PMID: 14615301

Nath AR, Fava EE, Beauchamp MS. 2011. Neural correlates of interindividual differences in children's audiovisual speech perception. *Journal of Neuroscience* **31**:13963–13971. DOI: https://doi.org/10.1523/JNEUROSCI.2605-11.2011, PMID: 21957257

Niemeyer W, Starlinger I. 1981. Do the blind hear better? investigations on auditory processing in congenital or early acquired blindness II. central functions. *International Journal of Audiology* **20**:510–515. DOI: https://doi.org/10.3109/00206098109072719

Nolte G, Ziehe A, Nikulin VV, Schlögl A, Krämer N, Brismar T, Müller KR. 2008. Robustly estimating the flow direction of information in complex physical systems. *Physical Review Letters* **100**:234101. DOI: https://doi.org/10.1103/PhysRevLett.100.234101, PMID: 18643502

Nolte G. 2003. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Physics in Medicine and Biology* **48**:3637–3652. DOI: https://doi.org/10.1088/0031-9155/48/22/002, PMID: 14680264

Obleser J, Eisner F, Kotz SA, a KS. 2008. Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *Journal of Neuroscience* **28**:8116–8123. DOI: https://doi.org/10.1523/JNEUROSCI.1290-08.2008, PMID: 18685036

Okada K, Rong F, Venezia J, Matchin W, Hsieh IH, Saberi K, Serences JT, Hickok G. 2010. Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. *Cerebral Cortex* **20**:2486–2495. DOI: https://doi.org/10.1093/cercor/bhp318, PMID: 20100898

Park H, Ince RA, Schyns PG, Thut G, Gross J. 2015. Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology* **25**:1649–1653. DOI: https://doi.org/10.1016/j.cub.2015.04.049, PMID: 26028433

Park H, Kayser C, Thut G, Gross J. 2016. Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *eLife* **5**:e14521. DOI: https://doi.org/10.7554/eLife.14521, PMID: 27146891

Patterson ML, Werker JF. 2003. Two-month-old infants match phonetic information in lips and voice. *Developmental Science* **6**:191–196. DOI: https://doi.org/10.1111/1467-7687.00271

Peelen MV, Bracci S, Lu X, He C, Caramazza A, Bi Y. 2013. Tool selectivity in left occipitotemporal cortex develops without vision. *Journal of Cognitive Neuroscience* **25**:1225–1234. DOI: https://doi.org/10.1162/jocn_a_00411, PMID: 23647514

Peelle JE, Davis MH. 2012. Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology* **3**:1–17. DOI: https://doi.org/10.3389/fpsyg.2012.00320, PMID: 22973251

Peelle JE, Gross J, Davis MH. 2013. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex* **23**:1378–1387. DOI: https://doi.org/10.1093/cercor/bhs118, PMID: 22610394

Peelle JE, Johnsrude IS, Davis MH. 2010. Hierarchical processing for speech in human auditory cortex and beyond. *Frontiers in Human Neuroscience* **4**:2486–2495. DOI: https://doi.org/10.3389/fnhum.2010.00051, PMID: 20661456

Pellegrino F, Coupé C, Marsico E. 2011. Across-language perspective on speech information rate. *Language* **87**:539–558. DOI: https://doi.org/10.1353/lan.2011.0057

Pietrini P, Furey ML, Ricciardi E, Gobbini MI, Wu WH, Cohen L, Guazzelli M, Haxby JV. 2004. Beyond sensory images: Object-based representation in the human ventral pathway. *PNAS* **101**:5658–5663. DOI: https://doi.org/10.1073/pnas.0400707101, PMID: 15064396

Poeppel D, Idsardi WJ, van Wassenhove V, Wassenhove van . 2008. Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**:1071–1086. DOI: https://doi.org/10.1098/rstb.2007.2160, PMID: 17890189

Poeppel D. 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication* **41**:245–255. DOI: https://doi.org/10.1016/S0167-6393(02)00107-3

Poirier C, Collignon O, Scheiber C, Renier L, Vanlierde A, Tranduy D, Veraart C, De Volder AG. 2006. Auditory motion perception activates visual motion areas in early blind subjects. *NeuroImage* **31**:279–285. DOI: https://doi.org/10.1016/j.neuroimage.2005.11.036, PMID: 16443376

Pons F, Lewkowicz DJ, Soto-Faraco S, Sebastián-Gallés N. 2009. Narrowing of intersensory speech perception in infancy. *PNAS* **106**:10598–10602. DOI: https://doi.org/10.1073/pnas.0904134106, PMID: 19541648

Reich L, Szwed M, Cohen L, Amedi A. 2011. A ventral visual stream reading center independent of visual experience. *Current Biology* **21**:363–368. DOI: https://doi.org/10.1016/j.cub.2011.01.040, PMID: 21333539

Ricciardi E, Vanello N, Sani L, Gentili C, Scilingo EP, Landini L, Guazzelli M, Bicchi A, Haxby JV, Pietrini P. 2007a. The effect of visual experience on the development of functional architecture in hMT+. *Cerebral Cortex* **17**: 2933–2939. DOI: https://doi.org/10.1093/cercor/bhm018, PMID: 17372275

Ricciardi E, Vanello N, Sani L, Gentili C, Scilingo EP, Landini L, Guazzelli M, Bicchi A, Haxby JV, Pietrini P. 2007b. The effect of visual experience on the development of functional architecture in hMT+. *Cerebral Cortex* **17**: 2933–2939. DOI: https://doi.org/10.1093/cercor/bhm018, PMID: 17372275

Rodd JM, Davis MH, Johnsrude IS. 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex* **15**:1261–1269. DOI: https://doi.org/10.1093/cercor/bhi009, PMID: 15635062

Rodd JM, Johnsrude IS, Davis MH. 2010. The role of domain-general frontal systems in language comprehension: evidence from dual-task interference and semantic ambiguity. *Brain and Language* **115**:182–188. DOI: https://doi.org/10.1016/j.bandl.2010.07.005, PMID: 20709385

Rosenblum LD, Schmuckler MA, Johnson JA. 1997. The McGurk effect in infants. *Perception & Psychophysics* **59**: 347–357. DOI: https://doi.org/10.3758/BF03211902, PMID: 9136265

Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ. 2007. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex* **17**:1147–1153. DOI: https://doi.org/10.1093/cercor/bhl024, PMID: 16785256

Röder B, Stock O, Bien S, Neville H, Rösler F. 2002. Speech processing activates visual cortex in congenitally blind humans. *European Journal of Neuroscience* **16**:930–936. DOI: https://doi.org/10.1046/j.1460-9568.2002.02147.x, PMID: 12372029

Sadato N, Pascual-Leone A, Grafman J, Ibañez V, Deiber MP, Dold G, Hallett M. 1996. Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* **380**:526–528. DOI: https://doi.org/10.1038/380526a0, PMID: 8606771

Schepers IM, Hipp JF, Schneider TR, Röder B, Engel AK. 2012. Functionally specific oscillatory activity correlates between visual and auditory cortex in the blind. *Brain* **135**:922–934. DOI: https://doi.org/10.1093/brain/aws014, PMID: 22366801

Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A. 2008. Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences* **12**:106–113. DOI: https://doi.org/10.1016/j.tics.2008.01.002, PMID: 18280772

Schroeder CE, Lakatos P. 2009. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences* **32**:9–18. DOI: https://doi.org/10.1016/j.tins.2008.09.012, PMID: 19012975

Schwartz JL, Savariaux C. 2014. No, there is no 150 ms lead of visual speech on auditory speech, but a range of audiovisual asynchronies varying from small audio lead to large audio lag. *PLoS Computational Biology* **10**: e1003743. DOI: https://doi.org/10.1371/journal.pcbi.1003743, PMID: 25079216

Scott SK, Blank CC, Rosen S, Wise RJ. 2000. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* **123**:2400–2406. DOI: https://doi.org/10.1093/brain/123.12.2400, PMID: 11099443

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. 1995. Speech recognition with primarily temporal cues. *Science* **270**:303–304. DOI: https://doi.org/10.1126/science.270.5234.303, PMID: 7569981

Smith SM, Nichols TE. 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage* **44**:83–98. DOI: https://doi.org/10.1016/j.neuroimage.2008.03.061, PMID: 18501637

Stephan KE, Penny WD, Moran RJ, den Ouden HEM, Daunizeau J, Friston KJ. 2010. Ten simple rules for dynamic causal modeling. *NeuroImage* **49**:3099–3109. DOI: https://doi.org/10.1016/j.neuroimage.2009.11.015

Striem-Amit E, Ovadia-Caro S, Caramazza A, Margulies DS, Villringer A, Amedi A. 2015. Functional connectivity of visual cortex in the blind follows retinotopic organization principles. *Brain* **138**:1679–1695. DOI: https://doi.org/10.1093/brain/awv083

Szymanski FD, Rabinowitz NC, Magri C, Panzeri S, Schnupp JW. 2011. The laminar and temporal structure of stimulus information in the phase of field potentials of auditory cortex. *Journal of Neuroscience* **31**:15787–15801. DOI: https://doi.org/10.1523/JNEUROSCI.1416-11.2011, PMID: 22049422

Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A. 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering* **44**:867–880. DOI: https://doi.org/10.1109/10.623056, PMID: 9282479

van Wassenhove V. 2013. Speech through ears and eyes: interfacing the senses with the supramodal brain. *Frontiers in Psychology* **4**:388. DOI: https://doi.org/10.3389/fpsyg.2013.00388, PMID: 23874309

Vouloumanos A, Druhen MJ, Hauser MD, Huizink AT. 2009. Five-month-old infants' identification of the sources of vocalizations. *PNAS* **106**:18867–18872. DOI: https://doi.org/10.1073/pnas.0906049106, PMID: 19846770

Weeks R, Horwitz B, Aziz-Sultan A, Tian B, Wessinger CM, Cohen LG, Hallett M, Rauschecker JP. 2000. A positron emission tomographic study of auditory localization in the congenitally blind. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* **20**:2664–2672. PMID: 10729347

Weikum WM, Vouloumanos A, Navarra J, Soto-Faraco S, Sebastián-Gallés N, Werker JF. 2007. Visual language discrimination in infancy. *Science* **316**:1159. DOI: https://doi.org/10.1126/science.1137686, PMID: 17525331

Zion Golumbic E, Cogan GB, Schroeder CE, Poeppel D. 2013a. Visual input enhances selective speech envelope tracking in auditory cortex at a "cocktail party". *Journal of Neuroscience* **33**:1417–1426. DOI: https://doi.org/10.1523/JNEUROSCI.3675-12.2013, PMID: 23345218

Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE. 2013b. Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron* **77**:980–991. DOI: https://doi.org/10.1016/j.neuron.2012.12.037, PMID: 23473326

Zoefel B, VanRullen R. 2015. Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *Journal of Neuroscience* **35**:1954–1964. DOI: https://doi.org/10.1523/JNEUROSCI.3484-14.2015, PMID: 25653354