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

Authors:

Markus J. van Ackeren[1]*, Francesca Barbero[2], Stefania Mattioni[1], Roberto Bottini[1], and Olivier Collignon[1,2]*

Affiliations:

[1] Center for Mind/Brain Studies, University of Trento, Trento, Italy

[2] Institute of research in Psychology (IPSY) & Institute of Neuroscience (IoNS) - University of Louvain (UCL), Louvain, Belgium

* Correspondence should be addressed to Olivier Collignon (olivier.collignon@uclouvain.be).

1   ## Abstract

2   The occipital cortex of early blind individuals (EB) activates during speech processing,

3   challenging the notion of a hard-wired neurobiology of language. But, at what stage of speech

4   processing do occipital regions participate in EB? Here we demonstrate that parieto-occipital

5   regions in EB enhance their synchronization to acoustic fluctuations in human speech in the theta-

6   range (corresponding to syllabic rate), irrespective of speech intelligibility. Crucially, enhanced

7   synchronization to the intelligibility of speech was selectively observed in primary visual cortex

8   in EB, suggesting that this region is at the interface between speech perception and

9   comprehension. Moreover, EB showed overall enhanced functional connectivity between

10   temporal and occipital cortices sensitive to speech intelligibility and altered directionality when

11   compared to the sighted group. These findings suggest that the occipital cortex of the blind

12   adopts an architecture allowing the tracking of speech material, and therefore does not fully

13   abstract from the reorganized sensory inputs it receives.

14

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 unknown how these two factors jointly 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, as well as its presence early in development are remarkable, and often used as an argument 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 "hardwired" and immutable to change by experience? Strong nativist proposals for linguistic innateness leave little room for plasticity due to experience (Bates, 1999), with for instance the proposition that we should conceive "the growth of language as analogous to the development of a bodily organ" (Chomsky, 1975, p. 11). However, studies in infants born with extensive damage to the typical brain's language areas may develop normal language abilities, demonstrating that the language network is subject to reorganization (Bates, 2005). Perhaps the most intriguing demonstration that the neurobiology of language is susceptible to change due to experience comes from studies in congenitally blind individuals showing functional selectivity to language

36   in primary and secondary 'visual' areas (Roder et al., 2002; Burton, 2003; Amedi et al., 2004;

37   Bedny et al., 2011; Arnaud et al., 2013). Such reorganization of the language network is

38   particularly fascinating since it arises in a system with no injuries of the core language network

39   (Bates, 2005; Atilgan et al., 2017).

40   However, the level of speech representation at which the occipital cortex of the early

41   blind participates remains poorly understood. Speech comprehension requires that the brain

42   extract meaning from the acoustic features of sounds (de Heer et al., 2017). Although several

43   neuroimaging studies have yielded valuable insights about the processing of speech in early

44   blind adults (Arnaud et al., 2013; Bedny et al., 2011; Burton, 2003; Lane et al., 2015; Röder et

45   al., 2002) and infants (Bedny et al., 2015), these methods do not adequately capture the fast and

46   continuous nature of speech processing. Because speech unfolds over time, understanding

47   spoken language rely on the ability to track the incoming acoustic signal in near real-time (Peelle

48   and Davis, 2012). Indeed, speech is a fluctuating acoustic signal that rhythmically excites

49   neuronal populations in the brain (Poeppel et al., 2008; Gross et al., 2013; Peelle et al., 2013).

50   Several studies have demonstrated that neuronal populations in auditory areas entrain to the

51   acoustic fluctuations that are present in human speech around the syllabic rate (Luo and Poeppel,

52   2007; Kayser et al., 2009; Szymanski et al., 2011; Zoefel and VanRullen, 2015). It has therefore

53   been suggested that entrainment reflects a key mechanism underlying hearing, for example, by

54   facilitating the parsing of individual syllables through adjusting the sensory gain relative to

55   fluctuations in the acoustic energy (Giraud and Poeppel, 2012; Peelle and Davis, 2012; Ding and

56   Simon, 2014). Crucially, because some regions that track the specific acoustic rhythm of speech

4

57   are sensitive to speech intelligibility, neural phase-locking is not solely driven by changes in the

58   acoustic cue of the auditory stimuli, but also reflects cortical encoding and processing of the

59   auditory signal (Peelle et al., 2013; Ding and Simon, 2014). Speech tracking is therefore an

60   invaluable tool to probe regions interfacing speech perception and comprehension (Poeppel et

61   al., 2008; Gross et al., 2013; Peelle et al., 2013).

62   Does the occipital cortex of early blind people synchronizes to speech rhythm? Is this

63   putative phase-locking of neural activity to speech influenced by its understandability?

64   Addressing those questions would provide novel insights onto the functional organization of

65   speech processing rooted in the occipital cortex of early blind people. In the current study, we

66   investigated if neuronal populations in blind occipital cortex synchronize to rhythmic dynamics

67   of speech. This tracking is demonstrated by relating the amplitude fluctuations in speech to

68   electromagnetic dynamics recorded from the participant's brain. To this end, we recorded the

69   brain activity of a group of early blind (EB; n=17) and sighted individuals (SI; n=16) with

70   Magnetoencephalography (MEG), while participants listened to short narrations from

71   audiobooks. If the occipital cortex of the blind entrains to speech rhythms, this will support the

72   idea that this region processes low-level acoustic features relevant for language understanding.

73   We further tested whether the putative phase locking of occipital responses in early blind people

74   benefits from linguistic information or only relates to acoustic information. To separate linguistic

75   and acoustic processes we relied on a noise-vocoding manipulation that spectrally distorted the

76   speech signal in order to gradually impair intelligibility but systematically preserve the slow

77   amplitude fluctuations responsible for speech rhythm (Shannon et al. 1995; Peelle et al., 2012).

78   Furthermore, going beyond differences in the local encoding of speech rhythm in the occipital

79   cortex, we also investigated whether early blindness alters the connectivity between occipital

80   and temporal regions sensitive to speech comprehension.

81

82   **Results**

83   **Story comprehension**

84   Participants listened to either natural speech segments (nat-condition) or to altered (using

85   vocoding) versions of these segments (see material and methods for details). In the 8-channel

86   vocoded condition, the voice of the speaker is highly distorted but intelligibility is unperturbed.

87   In contrast, in the 1-channel vocoded condition, speech is entirely unintelligible. After listening

88   to each speech segment, participants were provided with a short statement about the segment,

89   and asked to indicate whether the statement was true or false. Behavioural performance on these

90   comprehension statements was analysed using linear mixed-effects models with maximum

91   likelihood estimation. The method is a linear regression that takes into account dependencies in

92   the data, as present in repeated measures designs. Blindness and Intelligibility were included as

93   fixed effects, while subject was modelled as a random effect. Intelligibility was nested in subjects.

94   Intelligibility had a significant effect on story comprehension ($\chi(2)$=110.7, p<.001). The effect of

95   blindness and the interaction between intelligibility and blindness were non-significant

96   ($\chi(1)$=1.14, p=.286 and $\chi(2)$=.24, p=.889). Orthogonal contrasts demonstrated that speech

97   comprehension was stronger in the nat and 8-channel condition versus the 1-channel condition

98   (b=.78, $t(62)$=14.43, p<.001, r=.88). There was no difference between the nat and the 8-channel

6

99   condition (b=.02, *t*(62), 1.39, *p*=.17). Thus, speech intelligibility was reduced in the 1-channel,

100   but not the 8-channel vocoded condition (Fig 1D). The lack of effect for the factor blindness

101   suggests that there is no evidence for a potential difference in comprehension, attention or

102   motivation between groups. Descriptive statistics of the group, and condition means are

103   depicted in Table 1.

104

105   Table 1. *Proportion of correct responses for the story comprehension questions. Depicted are*

106   *the number of subjects (N) in each group, the mean (M), and the standard error of the mean (SE)*

107   *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

108

109   **Tracking intelligible speech in blind and sighted temporal cortex**

110   The overall coherence spectrum, which highlights the relationship between the amplitude

111   envelope of speech and the signal recorded from the brain, was maximal over temporal sensors

112   between 6-7 Hz (Fig 1E). This first analysis was performed on the combined dataset, and hence

113   does not cause the risk of a bias or circularity for subsequent analyses, targeting group

114   differences. The peak in the current study is slightly higher than what has been reported in

115   previous studies (Gross et al., 2013). A likely explanation for this shift is the difference in syllabic

116   rate between English (~6.2Hz), used in previous studies, and Italian (~7Hz) (Pellegrino et al.,

117   2011). The syllabic rate is the main carrier of amplitude fluctuations in speech, and thus most

118   prone to reset oscillatory activity.

119   To optimally capture the temporal scale of cerebro-acoustic coherence effects (6-7Hz),

120   and achieve a robust source estimate, source reconstruction was performed on two separate

121   frequency windows (6±2, and 7±2Hz) for each subject and condition. The two source images

122   were averaged for subsequent analysis, yielding a single image representing the frequency range

123   4-9Hz. The source-reconstructed coherence at the frequency of interest confirmed that cerebro-

124   acoustic coherence was strongest in bilateral temporal lobes including primary auditory cortex.

125   (Fig 1F).

126   To test whether envelope tracking in the current study is modulated by intelligibility, we

127   compared the coherence maps for the intelligible (nat), versus non-intelligible (1-channel)

128   condition, in all groups combined (i.e., EB and SI), with dependent-samples permutation t-tests

129   in SPM. The resulting statistical map (Fig 1G, *p*<.05, FWE-corrected), revealed a cluster in right

130   superior, and middle temporal cortex (STG, MTG) where synchronization to the envelope of

131   speech was stronger when participants had access to the content of the story.

132 Table 2. *MNI coordinates and p-values for the three contrasts tested with 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

133

134 **Speech tracking in blind parieto-occipital cortex**

135 Whether early blind individuals recruit additional neural substrate for tracking the

136 envelope of speech was tested using independent-samples permutation t-tests in SPM

137 contrasting coherence maps between EB and SI for all three conditions combined. The statistical

138 maps (Fig. 1H, *p*<.05, FWE-corrected) revealed enhanced coherence in EB versus SI in parieto-

139 occipital cortex along the medial wall, centered on bilateral Precuneus, branching more

140 extensively into the right hemisphere (see Table 2). These results demonstrate that neuronal

141 populations in blind parieto-occipital cortex synchronize to the temporal dynamics in speech.

142 Finally, to investigate if, and where envelope tracking is sensitive to top-down predictions

143 during intelligible speech comprehension, we subtracted the unintelligible (1-channel) from the

144 intelligible (nat) condition, and computed independent-samples t-tests between groups. As

145 highlighted in the introduction, sensitivity to language is most surprising in early 'visual' areas of

9

146   blind occipital cortex. Especially the role of the area around Calcarine sulcus (V1) is currently

147   debated (Burton et al., 2002; Roder et al., 2002; Amedi et al., 2003). As there is no prior on the

148   potential effect size for this contrast we restricted the analysis to the area around primary visual

149   cortex. The search volume was constructed from four 10mm spheres around coordinates in

150   bilateral Calcarine sulcus ([-7, -81, -3]; [-6, -85, 4]; [12, -87, 0]; [10, -84, 6]). The coordinates were

151   extracted from a previous study on speech comprehension in EB (Burton et al., 2002). The

152   resulting mask also extends into regions highlighted in similar studies by other groups (Roder et

153   al., 2002; Bedny et al., 2011).

154   A significant effect for the interaction between blindness and intelligibility [(nat$_{EB}$ – 1-

155   channel$_{EB}$) - (nat$_{SI}$ – 1-channel$_{SI}$)] was observed in right Calcarine sulcus. To explore the spatial

156   specificity of the effect, we show the whole-brain statistical map for the interaction between

157   intelligibility and blindness at a more liberal threshold (p<.005, uncorrected) in Fig 1I. The figure

158   shows that intelligible speech selectively engages the area around the right calcarine sulcus in

159   the blind versus sighted. Specifically, the region corresponding to right 'visual' showed enhanced

160   sensitivity to intelligible speech in EB versus SI (p<.05, FWE-corrected). The analysis contrasting

161   the two intelligible conditions (nat, 8-channel), did however not yield a significant effect in the

162   EB, suggesting that the effect is not likely driven by the low-level degrading of the stimulus alone,

163   but rather the intelligibility of the speech segment. Follow-up post-hoc comparisons between

164   the two groups revealed that coherence was stronger during the intelligible condition for EB

165   versus SI (t(30)=3.09, p=.004), but not during the unintelligible condition (t(30)=-1.08, p=.29).

166   Nevertheless, inspection of the data suggests that the overall interaction effect was influenced

10

167   by high coherence during the unintelligible condition in the SI group. For illustration purposes,

168   cerebro-acoustic coherence from functional peak locations for intelligibility (across groups) in

169   STG, blindness in POC (Fig 1G), and the interaction, are represented as boxplots in Fig 1J.

170   **Changes in occipito-temporal connectivity in the early blind**

171   To further investigate whether cerebro-acoustic peaks in CS, and STG, which are also

172   sensitive to intelligible speech in EB, are indicative of a more general re-organization of the

173   network, we applied functional connectivity analysis. Statistical analysis of the connectivity

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

175   Using linear mixed-effects models, blindness (EB, SI), intelligibility (nat,1-channel), and

176   the interaction between blindness and intelligibility were added to the model in a hierarchical

177   fashion. Blindness and intelligibility were modelled as fixed effects, while subject was a random

178   effect. Intelligibility was nested in subject. A main effect was observed only for blindness

179   ($\chi(1)$=4.32, p=.038). That is, connectivity was overall stronger for EB versus SI (Fig 2 A-B). The

180   main effect of intelligibility, and the interaction between intelligibility and blindness were non-

181   significant (p=.241, and p=.716 respectively)

182   Subsequently, linear mixed-effects models were applied to test for the effects blindness,

183   and intelligibility on the directional information flow (PSI) between CS and STG. The fixed and

184   random effects structure was the same as described in the previous analysis. Here, only the main

185   effect of blindness was significant ($\chi(1)$=4.54, p=.033). That is, the directional information flow

186   differs between groups. The effect of intelligibility and the interaction between intelligibility and

187   blindness were both non-significant (p=.51, p=.377). Follow-up post-hoc one-sample t-tests on

188   the phase slope estimates for each group individually revealed a significant direction bias from

189   CS to STG for SI ($t(14)$=-2.22, p=.044). No directional bias was found for EB ($t(16)$=.74, p=.47) As

190   depicted in Fig 2 C-D, these results suggest that CS projects predominantly to STG in SI, whereas

191   in EB this interaction is likely more balanced, and trending in the opposite direction.

192   **Discussion**

193   We took advantage on the statistical and conceptual power offered by correlating

194   magnetoencephalographic recordings with the envelope of naturalistic continuous speech to

195   test if neuronal populations in the early blind's occipital cortex synchronize to the temporal

196   dynamics present in human speech. Using source localized MEG activity, we confirm and extend

197   previous research (Scott et al., 2000; Davis and Johnsrude, 2003; Narain et al., 2003; Rodd et al.,

198   2005, 2010; Okada et al., 2010; Peelle et al., 2010) by showing that temporal brain regions

199   entrain to the natural rhythm of speech signal in both blind and sighted groups (Fig 1F).

200   Strikingly, we show that following early visual deprivation occipito-parietal regions enhance their

201   synchronization to the acoustic rhythm of speech (Fig 1G), independently of language content.

202   That is, the region around the Precuneus/Cuneus area, seems to be involved in low-level

203   processing of acoustic information alone. Previous studies have demonstrated that this area

204   enhance its response to auditory information after early but not late visual deprivation (Collignon

205   et al., 2013) and maintain such elevated response to sound even after sight-restoration early in

206   life (Collignon et al., 2015). These studies did not find that this crossmodal response was related

207   to a specific cognitive function but rather pointed to a more general response to sounds

208   (Collignon et al., 2013, 2015). Interestingly, the current data suggest that this area is capable of

209   parsing complex temporal patterns in sounds but is however insensitive to the intelligibility of

210   speech, again suggesting a more general role in sound processing.

211   In order to isolate the brain regions modulated by speech comprehension, we identified

212   those regions showing a greater activity for amplitude modulations that convey speech

213   information compared to amplitude modulations that do not. In addition to the right STG

214   observed in both groups (Fig 1H), the blind showed enhanced cerebro-acoustic coherence

215   during intelligible speech at the lowest level of the occipital hierarchy, in the vicinity of calcarine

216   sulcus (V1; Fig. 1I). This pattern of local encoding was accompanied by enhanced occipito-

217   temporal connectivity during speech comprehension in EB compared to SI. While SI show the

218   expected feed-forward projections from occipital to temporal regions (Lamme et al., 1998), EB

219   show a more balanced connectivity profile, trending towards the reverse temporal to occipital

220   direction. These findings support the idea of a reverse hierarchical model (Büchel, 2003) of the

221   occipital cortex in EB where the regions typically coding for "low-level" visual features in the

222   sighted (e.g. visual contrast or orientation) participate in higher-level function (e.g. speech

223   intelligibility) in EB. Indeed, previous studies have found increased activity in the primary "visual"

224   cortex of early blind people during Braille reading (Sadato et al., 1996; Burton et al., 2002, 2012),

225   verbal memory and verb generation tasks (Amedi et al., 2003) and during auditory language-

226   related processing(Bedny et al., 2011). In line with our results, activity in primary occipital regions

227   in early blinds is stronger in a semantic versus a phonologic task (Burton et al., 2003) and vary as

228   a function of syntactic and semantic complexity (Roder et al., 2002; Bedny et al., 2011; Lane et

229   al., 2015). Moreover, repetitive transcranial magnetic stimulation (rTMS) over the occipital pole

13

230   induces more semantic than phonologic errors in a verb-generation task in early blind people

231   (Amedi 2003). However, our results however clearly suggest that the involvement of the occipital

232   pole for language is not fully abstracted from sensory inputs as previously suggested (Bedny,

233   2017) since we show that occipital regions entrain to the envelope of speech and is enhanced

234   by its intelligibility. Occipital participation in tracking the flow of speech in EB may constitute an

235   adaptive strategy to boost perceptual sensitivity at informational peaks in language.

236   **Mechanistic origins of envelope tracking in occipital cortex**

237   Although language is not the only cognitive process that selectively activate the occipital

238   cortex of early blind people, it is arguably one of the most puzzling. The reason is that

239   reorganization of other processes such auditory motion perception, and tactile object

240   recognition appears to follow the topography of the functionally equivalent visual processes in

241   the sighted brain (Ricciardi et al., 2007a; Amedi et al., 2010; Dormal et al., 2016). For example,

242   the hMT+/V5 complex, typically involved in visual motion in the sighted, selectively processes

243   auditory (Poirier et al., 2006; Dormal et al., 2016; Jiang et al., 2016) or tactile (Ricciardi et al.,

244   2007b) motion in the blind. However, in the case of language, such recruitment is striking in light

245   of the cognitive and evolutionary differences between vision and language (Bedny et al., 2011).

246   This lead to the proposal that at birth, human cortical areas are cognitively pluripotent: capable

247   of assuming a broad range of unrelated cognitive functions(Bedny, 2017). However, this

248   argument resides on the presupposition that language has no computational relation with vision.

249   But does this proposition reflect with what we know about the relation between the visual system

250   and the classical language network? Rhythmic information in speech has a well-known language-

251   related surrogate in the visual domain, with the observation of lip movements (Lewkowicz and

252   Hansen-Tift, 2012; Park et al., 2016). Indeed, both acoustic and visual speech signals exhibit

253   rhythmic temporal patterns at prosodic and syllabic rates (Caplier et al., 2009; Schwartz and

254   Savariaux, 2014; Giordano et al., 2017). The perception of lip kinematics that are naturally linked

255   to amplitude fluctuations in speech serves as an important vehicle for the everyday use of

256   language (Kuhl and Meltzoff, 1982; Weikum et al., 2007; Lewkowicz and Hansen-Tift, 2012) and

257   help language understanding, especially in noisy conditions (Ross et al., 2007). Actually, reading

258   lips in the absence of any sound activates both primary and association auditory regions

259   overlapping with regions active during the actual perception of spoken words(Calvert et al.,

260   1997). The synchronicity between auditory and visual speech entrain rhythmic activity in the

261   observer's primary auditory and visual regions and facilitate perception by aligning neural

262   excitability with acoustic or visual speech features (Schroeder et al., 2008; Schroeder and Lakatos,

263   2009; Giraud and Poeppel, 2012; Mesgarani and Chang, 2012; Peelle and Davis, 2012; van

264   Wassenhove, 2013; Zion Golumbic et al., 2013b; Park et al., 2016; Giordano et al., 2017). These

265   results strongly suggest that both auditory and visual components of speech are processed

266   together at the earliest level possible in neural circuitry based on the shared slow temporal

267   modulations (around 2–7 Hz range) present across modalities (Caplier et al., 2009).

268   Corroborating this idea, it has been demonstrated that neuronal populations in visual cortex

269   follow the temporal dynamics of lip movements in sighted individuals, similar to the way temporal

270   regions follow the acoustic and visual fluctuations of speech (Luo et al., 2010; Zion Golumbic et

271   al., 2013a; Park et al., 2016). Like the temporal cortex, occipital cortex in the sighted also shows

272   enhanced lip tracking when attention is directed to speech content. This result highlights the fact

15

273   that a basic oscillatory architecture for tracking the dynamic aspects of (visual-) speech in occipital

274   cortex exists even in sighted individuals.

275   Importantly, audiovisual integration of the temporal dynamics of speech has been

276   suggested to play a key role when learning speech early in life: young infants detect, match and

277   integrate the auditory and visual temporal coherence of speech (Kuhl and Meltzoff, 1982;

278   Rosenblum et al., 1997; Lewkowicz, 2000, 2010; Brookes et al., 2001; Patterson and Werker,

279   2003; Lewkowicz and Ghazanfar, 2006; Kushnerenko et al., 2008; Bristow et al., 2009; Pons et

280   al., 2009; Vouloumanos et al., 2009; Lewkowicz et al., 2010; Nath et al., 2011). For instance,

281   young infants between 4 and 6 months of age can detect their native language from lip

282   movements only (Weikum et al., 2007). Around the same period, children detect synchrony

283   between lip movements and speech sounds, and distribute more attention towards the mouth

284   than towards the eyes (Lewkowicz and Hansen-Tift, 2012). Linking what they hear to the lip

285   movements they observe may provide young infants with a stepping-stone towards language

286   production (Lewkowicz and Hansen-Tift, 2012). Moreover, infants aged 10 weeks already exhibit

287   a McGurk effect; highlighting again the early multisensory nature of speech perception

288   (Rosenblum et al., 1997). Taken together, these results suggest that an audio-visual link between

289   observing lip movements and hearing speech sounds is present at very early developmental

290   stages, potentially from birth, which helps infants acquire their first language.

291   In the context of these evidences, the current results may support the biased connectivity

292   hypothesis of cross-modal reorganization (Reich et al., 2011; Collignon et al., 2012; Hannagan

293   et al., 2015; Striem-amit et al., 2015). Indeed, it has been argued that reorganization in blind

16

294   occipital cortex might be constrained by functional pathways to other sensory- and cognitive

295   systems that are also present in sighted individuals (Hannagan; Johnson). This hypothesis may

296   explain the overlap in functional specialization between blind and sighted individuals (Collignon

297   et al., 2011; Dormal et al., 2016; He et al., 2013; Jiang et al., 2016; Peelen et al., 2013; Pietrini

298   et al., 2004; Poirier et al., 2006). In the context of our experiment, sensitivity to acoustic dynamics

299   of intelligible speech in blind occipital cortex could arise from pre-existing occipito-temporal

300   pathways connecting the auditory and visual system that are particularly important for the early

301   developmental stages of language acquisition. In fact, the reorganization of these potentially

302   pre-disposed pathway to process language content would explain how language-selective

303   response may appear as early as in 3yo blind children (Bedny et al., 2015).

304   Previous studies have suggested that language processing in the occipital cortex arise

305   through top-down projections from frontal region typically associated to the classical language

306   network (Bedny et al., 2011; Deen et al., 2015) and that the representational content is symbolic

307   and abstract, rather than sensory(Bedny, 2017). Our results contrasts with this view by showing

308   that neuronal populations in (peri-) calcarine cortex align to the temporal dynamics of intelligible

309   speech, and are functionally connected to areas sensitive to auditory information in temporal

310   cortex. In sighted individuals, regions of the temporal lobe including STS are sensitive to acoustic

311   features of speech, whereas higher-level regions such as anterior temporal cortex and left inferior

312   frontal gyrus are relatively insensitive to these features and therefore do not entrain to the syllabic

313   rate of speech (Davis and Johnsrude, 2003; Hickok and Poeppel, 2007; see confirmation in figure

314   1F-G). This suggests that occipital areas respond to speech at a much lower – sensory – level

17

315   than previously thought in EB, potentially based on the reorganization of existing multisensory

316   pathways connecting 'auditory' and 'visual' centres in the brain. Functional dependencies

317   between sensory systems indeed exist between the earliest stages of sensory processing in

318   humans (Ghazanfar and Schroeder, 2006; Kayser et al., 2008; Schroeder and Lakatos, 2009;

319   Murray et al., 2016) and in nonhuman primates (Falchier et al., 2002; Lakatos et al., 2007;

320   Schroeder and Lakatos, 2009)(Lakatos et al. 2005, 2007, suggesting an evolutionarily conserved

321   mechanism of sensory processing. We therefore postulate that brain networks typically

322   dedicated to the integration of audio-visual speech signal might be reorganized in the absence

323   of visual inputs and lead to an extension (eg. lack of synaptic pruning) of speech tracking in the

324   occipital cortex. Building on this connectivity bias, the occipital pole may extend its sensitivity to

325   the intelligibility of speech, a computation this region is obviously not dedicated for in the first

326   place.

327   Actually, a number of recent studies have suggested that visual deprivation reinforces the

328   functional connections between the occipital cortex and auditory regions typically classified as

329   the language network (Hasson et al., 2016; Schepers et al., 2012). Previous studies relying on

330   dynamic causal modelling bring partial support to the idea that auditory information reaches the

331   reorganized occipital cortex of the blind through direct temporo-occipital connection rather than

332   using subcortical (Klinge et al., 2010) or top-down pathways (Collignon et al., 2013). In support

333   of these studies, we observed that the overall magnitude of functional connectivity between

334   occipital and temporal cortex is higher in the blind versus sighted individuals during natural

335   speech comprehension. Moreover, directional connectivity analysis revealed that the interaction

336   between the two cortices is also qualitatively different; sighted individuals show a strong feed-

337   forward drive towards temporal cortex, while blind individuals show a more balanced information

338   flow, and a trend in the reverse direction. Our hypothesis that speech processing in the occipital

339   cortex of EB arises from direct occipito-temporal connection is consistent with previous results

340   showing that entrainment of theta activity in auditory cortex is insensitive to top-down influences

341   from left inferior frontal regions (Kayser et al., 2015; Park et al., 2015).

342   **Speech comprehension in the right hemisphere**

343   We observed that neuronal populations in right superior temporal cortex synchronize to

344   the temporal dynamics of intelligible, rather than non-intelligible speech in both groups

345   combined. Why does speech intelligibility modulate temporal regions of the right, and not the

346   left, hemisphere? According to an influential model in speech comprehension, the asymmetric

347   sampling in time model (AST; Giraud and Poeppel, 2012; Hickok and Poeppel, 2007; Poeppel,

348   2003), there is a division of labour between left- and right auditory cortices (Poeppel, 2003;

349   Boemio et al., 2005; Hickok and Poeppel, 2007), with left auditory cortex being more sensitive

350   to high frequency information (+20Hz), whereas right temporal cortex responds mostly to low

351   frequency information (~6Hz), such as syllable sampling and prosody (Belin et al., 1998; Poeppel,

352   2003; Boemio et al., 2005; Obleser et al., 2008; Giraud and Poeppel, 2012). Several studies have

353   shown that the right hemisphere more specifically involves in the representation of connected

354   speech (Bourguignon et al., 2013; Fonteneau et al., 2015; Horowitz-Kraus et al., 2015;

355   Alexandrou et al., 2017), and previous studies directly demonstrated the prevalence of speech-

356   to-brain entrainment in delta and theta bands in the right hemisphere more than the left during

357   the listening of sentences or stories (Luo and Poeppel, 2007; Abrams et al., 2008; Gross et al.,

358   2013; Giordano et al., 2017) The present study therefore replicates those results by showing

359   enhanced phase coupling between the right hemisphere and the speech envelope at the syllabic

360   rate (low-frequency phase of speech envelope), consistent with the AST-model. An interesting

361   observation in the current study is that right hemispheric sensitivity to intelligible speech in

362   temporal areas coincides with the enhanced right hemispheric sensitivity to intelligible speech

363   in blind occipital cortex. An important open question for future research concerns the relative

364   behavioral contribution of occipital and perisylvian cortex to speech understanding.

365

Conclusion

367   We demonstrate that the right primary "visual" cortex synchronizes to the temporal

368   dynamics of intelligible speech, at the rate of syllable transitions in language (~6-7Hz). These

369   results demonstrate that the involvement of this neural population relates to the sensory signal

370   of speech and therefore contrasts with the proposition that occipital involvement in speech

371   processing is abstracted from its sensory input and purely reflect higher-level operations similar

372   to those observed in prefrontal regions(Bedny, 2017). Blindness, due to the absence of

373   organizing visual input, leaves the room open for sensory and functional colonization of occipital

374   regions. This colonization will however not be stochastic, but will be constrained by modes of

375   information processing natively anchored in specific brain regions and networks. Even if the exact

376   processing mode has still to be unveiled by future research, we presume that language map

377   onto occipital cortex building on pre-existing oscillatory architecture typically linking auditory

20

378   and visual speech rhythm (Caplier et al., 2009). Our study therefore supports the view that the

379   development of functional specialisation in the human cortex is the product of a dynamic

380   interplay between genetic and environmental factors during development, rather than being

381   predetermined at birth (Elman et al., 1996).

382

383   **Material and Methods**

384   **Participants**

385   Seventeen early blind (11 female, mean ± SD, 32.9 ± 10.19 years, range, 20-67 years) and sixteen

386   sighted individuals (10 female, mean ± SD, 32.2 ± 9.92 years, range, 20-53 years) participated

387   in the current study. There was no age difference between the blind and the sighted group (t(30)

388   = .19, p=.85). All participants were either totally blind, or severely visually impaired from birth,

389   however, two participants reported residual visual perception before the age of 3, one before

390   the age of 4 and one participant lost sight completely at age 10. Causes of vision loss were

391   damage to, or detached retina (10), damage to the optic nerve (3), infection of the eyes (1),

392   microphtalmia (2), and hypoxia (1). While some participants reported residual light perception,

393   none was able to use vision functionally. All participants were proficient braille readers, and

394   native speakers of Italian. None of them suffered from a known neurological or peripheral

395   auditory disease. The data of one sighted individual was not considered due to the discovery of

396   a brain structural abnormality unknown to the experimenters at the time of the experiment. The

397   project was approved by the local ethical committee at the University of Trento. In agreement

21

398   with the Declaration of Helsinki, all participants provided written informed consent to participate

399   in the study.

400

Experimental design

402   Auditory stimuli were delivered into the magnetically shielded MEG room via stereo

403   loudspeakers using a Panphonics Sound Shower 2 amplifier at a comfortable sound level, which

404   was the same for all participants. Stimulus presentation was controlled via the Matlab

405   Psychophysics Toolbox 3 [http://psychtoolbox.org] running on a Dell Alienware Aurora PC under

406   Windows 7 (64 bit). Both sighted and blind participants were blindfolded during the experiment,

407   and the room was dimly lit to allow for visual monitoring of the participant via a video feed from

408   the MEG room. Instructions were provided throughout the experiment using previous recordings

409   from one of the experimenters (FB).

410   The stimuli consisted of 14 short segments (~1min) from popular audiobooks (e.g., Pippi

411   Longstocking and Candide) in Italian (nat). Furthermore, channel-vocoding in the Praat software

412   was used to produce two additional control conditions. First, the original sound file was band-

413   pass filtered into 1 (1-channel) or 8 (8-channel) logarithmically spaced frequency bands. The

414   envelope for each of these bands was computed, filled with Gaussian white noise, and the

415   different signals were recombined into a single sound file. The resulting signal has an amplitude

416   envelope close to the original, while the fine spectral detail is gradually distorted. Perceptually,

417   the voice of the speaker is highly distorted in the 8-channel condition, however intelligibility is

418   unperturbed. In contrast, in the 1-channel condition, speech is entirely unintelligible. In total, 42

22

419   sound files were presented in a pseudo-randomized fashion, distributed among seven blocks.

420   Each block contained two sound files from each condition, and the same story was never used

421   twice in the same block. Examples of the stimuli are provided in supplementary media content

422   (S1).

423   To verify story comprehension, each speech segment was followed by a single-sentence

424   statement about the story. Participants were instructed to listen to the story carefully, and judge

425   whether the statement at the end was true or false, using a nonmagnetic button box. Responses

426   were provided with the index and middle finger of the right hand.

427

428   **MEG data acquisition and preprocessing**

429   MEG was recorded continuously from a 306 triple sensor (204 planar gradiometers; 102

430   magnetometers) whole-head system (Elekta Neuromag, Helsinki, Finland) using a sampling rate

431   of 1kHz and online band-bass filter between 0.1 and 300Hz. The headshape of each individual

432   participant was measured using a Polhemus FASTRAK 3D digitizer. Head position of the subject

433   was recorded continuously using five localization coils (forehead, mastoids).

434   Data pre-processing was performed using the open-source Matlab toolbox Fieldtrip

435   [www.fieldtriptoolbox.org], as well as custom code. First the continuous data was filtered (high-

436   pass Butterworth filter at 1Hz, low-pass Butterworth filter at 170-Hz, and DFT filter at 50,100, and

437   150Hz to remove line-noise artefacts in the signal), and downsampled to 256Hz. Next, the data

438   were epoched into segments of 1 second for subsequent analysis.

439   Rejection of trials containing artefacts and bad channels was performed using a semi-

440   automatic procedure. First, outliers were rejected using a pre-screening based on the variance

441   and range in each trial/channel. Then, algorithmically guided visual inspection of the raw data

442   was performed to remove any remaining sources of noise.

443

444   **Extraction of the speech amplitude envelope**

445   Amplitude envelopes of the stories were computed using the Chimera toolbox (Caplier et al.,

446   2009) and custom code, following the procedure described by Gross and colleagues (Gross et

447   al., 2013). First, the sound files were band-pass filtered between 100 and 1000Hz into 9

448   frequency-bands, using a 4th order Butterworth filter. The filter was applied in forward and

449   backward direction to avoid any phase shifts with respect to the original signal, and frequency-

450   bands were spaced with equal width along the human basilar membrane. Subsequently, the

451   analytic amplitude for each filtered segment was computed as the absolute of the Hilbert

452   transform. Finally, the amplitude envelopes of all 9 bands were summed and scaled to a

453   maximum value of 1. The resulting envelope was combined with the MEG data, and processed

454   identically henceforth.

455

456   **Analysis of cerebro-acoustic coherence**

457   To determine where, and at what temporal scale neuronal populations follow the temporal

458   dynamics of speech, we computed spectral coherence between the speech envelope and the

459   MEG signal. Coherence is a statistic that quantifies the phase relationship between two signals,

460   and can be used to relate oscillatory activity in the brain with a peripheral measure such as a

461   speech signal (Peelle et al., 2013). The first analysis was conducted in sensor space, across

462   conditions and participants, to determine at what temporal scale coherence between the speech

463   envelope and the MEG signal is strongest. To this end, a Hanning taper was applied to the 1s

464   data segments, and Fourier transformation was used to compute the cross-spectral density

465   between 1 and 30Hz, with a step size of 1Hz. To render the coherence values more normally

466   distributed, a Fisher z-transform was applied by computing the inverse hyperbolic tangent

467   (atanh).

468   Source-space analysis was centred on the coherence frequency-band of interest (FOI),

469   identified in the sensor space analysis. Source reconstruction was performed using a frequency

470   domain beamformer called Dynamic Imaging of Coherent Sources (DICS) (Gross et al., 2001;

471   Liljeström et al., 2005). DICS was used to compute cerebro-acoustic coherence for all locations

472   on a 3-dimensional grid (8x8x8mm). The forward model was based on a realistic single-shell

473   headmodel(Nolte, 2003) for each individual. As structural scans could not be acquired for all

474   participants, we approximated individual anatomy by warping a structural MNI template brain

475   (MNI, Montreal, Quebec, Canada; www.bic.mni.mcgill.ca/brainweb) into individual headspace

476   using the information from each participant's headshape.

477

478   **Source connectivity analysis**

479   Functional connectivity between occipital and temporal cortex was computed by

480   extracting virtual sensor time-series at the locations of interest in CS and STG using time-domain

481   beamforming. These virtual sensor time series were used to compute non-directional and

482   directional connectivity metrics.

483   Virtual sensor time-series at the two locations of interest were extracted using a time-

484   domain vector-beamforming technique called linear constrained minimum variance (LCMV)

485   beamforming (Van Veen et al., 1997). First, average covariance matrices were computed for each

486   participant to estimate common spatial filter coefficients. These filter coefficients were multiplied

487   with the single-trial cleaned MEG data. To reduce the resulting three-dimensional time-series to

488   one, singular value decomposition (SVD) was applied, resulting in a single time-series for each

489   trial and participant.

490   Functional connectivity between the two regions was computed at peak locations using

491   the phase-locking value (PLV) (Lachaux et al., 1999). First, single-trial virtual sensor time courses

492   were converted to the frequency domain (0-50Hz) using Fourier transformation. The data was

493   padded to 2s and a Hanning taper was applied to reduce spectral leakage. Coherence was used

494   as a proxy for functional connectivity. To disentangle phase consistency between regions from

495   joint fluctuations in power, the spectra were normalized with respect to the amplitude, resulting

496   in an estimate of the phase-locking (Lachaux et al., 1999) between regions. Connectivity

497   estimates were normalized using a Fischer z-transform (atanh), as in the analysis of cerebro-

498   acoustic coherence.

499   PLV is a symmetric proxy for connectivity and does not allow for inferences regarding the

500   direction of information flow between two regions of interest. To test whether differences in

501   functional connectivity between EB and SI are accompanied by changes in the directionality of

502   the flow of information between the regions of interest, we computed the phase slope index

503   (PSI) (Nolte et al., 2008). The phase slope index deduces net directionality from the time delay

504   between two time series (x1 and x2). Timeseries x1 is said to precede, and hence drive timeseries

505   x2, in a given frequency band, if the phase difference between x1 and x2 increases with higher

506   frequencies. Consequently, a negative phase slope reflects a net information flux in the reverse

507   direction, that is, from x2 to x1. Here, we computed the phase slope index using a bandwidth of

508   ±5Hz around the frequencies of interest. Following the recommendation by Nolte and

509   colleagues (Nolte et al., 2008), PSI estimates were normalized with the standard error, which was

510   computed using the jackknife method.

511

512   **Statistical analysis**

513   Statistical testing of the behavioural comprehension scores as well as the connectivity

514   estimates was performed using linear mixed-effects models in R. Differences between conditions

515   and groups in source space were evaluated using Statistical Parametric Mapping (SPM12), and

516   the Threshold-Free Cluster Enhancement (TFCE) toolboxes in Matlab. TFCE (Smith and Nichols,

517   2009) computes new values for each voxel in a statistical map as a function of the original voxel

518   value and the values of the surrounding voxels. By enhancing the statistical values in voxels with

519   a high T-value that are also part of a local cluster, TFCE optimally combines the benefits of voxel-

520   wise and cluster-based methods. TFCE was applied to the whole-brain contrasts. Final correction

521   for multiple comparisons was applied using a maximum statistic based on 1000 permutations of

522   group membership (independent testing) or conditions (dependent testing). Here, we applied a

523   variance smoothing of 15mm FWHM to reduce the effects of high spatial frequency noise in the

524   statistical maps. The smoothing kernel used in the current study is higher than in comparable

525   fMRI studies due to the inherent smoothness of the source-reconstructed MEG data.

526

27

28

538 **References**

539 Abrams D a, Nicol T, Zecker S, Kraus N (2008) Right-Hemisphere Auditory Cortex Is Dominant

540   for Coding Syllable Patterns in Speech. J Neurosci 28:3958–3965.

541 Alexandrou AM, Saarinen T, Mkel S, Kujala J, Salmelin R (2017) The right hemisphere is

542   highlighted in connected natural speech production and perception. Neuroimage 152:628–

543   638.

544 Amedi A, Floel A, Knecht S, Zohary E, Cohen LG (2004) Transcranial magnetic stimulation of the

545   occipital pole interferes with verbal processing in blind subjects. NatNeurosci 7:1266–1270

546 Amedi A, Raz N, Azulay H, Malach R, Zohary E (2010) Cortical activity during tactile exploration

547   of objects in blind and sighted humans. 28:143–156.

548 Amedi A, Raz N, Pianka P, Malach R, Zohary E (2003) Early "visual" cortex activation correlates

549   with superior verbal memory performance in the blind. Nat Neurosci 6:758–766.

550 Arnaud L, Sato M, Ménard L, Gracco VL (2013) Repetition suppression for speech processing in

551   the associative occipital and parietal cortex of congenitally blind adults. PLoS One 8:e64553

552 Atilgan H, Collignon O, Hasson U (2017) Structural neuroplasticity of the superior temporal plane

553   in early and late blindness. Brain Lang 170:71–81.

554 Bates E (1999) Language and the infant brain. J Commun Disord 32:195–205.

555 Bates ED of PU of CASDLJC a US (2005) Plasticity, Localization, and Language Development.

556   Milbrath, Constance; Ed; Biol Knowl Revisit From Neurogenes to Psychogenes 9351:205–

557   253.

558 Bedny M (2017) Evidence from Blindness for a Cognitively Pluripotent Cortex. Trends Cogn Sci

559   21:637–648 Available at: http://dx.doi.org/10.1016/j.tics.2017.06.003.

560   Bedny M, Pascual-Leone A, Dodell-Feder D, Fedorenko E, Saxe R (2011) Language processing

561   in the occipital cortex of congenitally blind adults. Proc Natl Acad Sci U S A 108:4429–4434

562   Bedny M, Richardson H, Saxe R (2015) "Visual" Cortex Responds to Spoken Language in Blind

563   Children. J Neurosci 35:11674–11681.

564   Belin P, Zilbovicius M, Crozier S, Thivard L, Fontaine A, Masure M-C, Samson Y (1998)

565   Lateralization of speech and auditory temporal processing. J Cogn Neurosci 10:536–540.

566   Berwick RC, Friederici AD, Chomsky N, Bolhuis JJ (2013) Evolution, brain, and the nature of

567   language. Trends Cogn Sci 17:98.

568   Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity

569   in human auditory cortices. Nat Neurosci 8:389–395.

570   Bourguignon M, De Tiège X, De Beeck MO, Ligot N, Paquier P, Van Bogaert P, Goldman S, Hari

571   R, Jousmäki V (2013) The pace of prosodic phrasing couples the listener's cortex to the

572   reader's voice. Hum Brain Mapp 34:314–326.

573   Bristow D, Dehaene-Lambertz G, Mattout J, Soares C, Gliga T, Baillet S, Mangin J-F (2009)

574   Hearing faces: how the infant brain matches the face it sees with the speech it hears. J Cogn

575   Neurosci 21:905–921.

576   Brookes H, Slater A, Quinn PC, Lewkowicz DJ, Hayes R, Brown E (2001) Three-Month-Old Infants

577   Learn Arbitrary Auditory-Visual Pairings between Voices and Faces. Infant Child Dev 10:75–

578   82.

579   Büchel C (2003) Cortical hierarchy turned on its head. Nat Neurosci 6:657–658

580   Burton H (2003) Visual cortex activity in early and late blind people. J Neurosci 23:4005–4011.

581   Burton H, Diamond JB, Mcdermott KB (2003) Dissociating cortical regions activated by semantic

582   and phonological tasks: a FMRI study in blind and sighted people. JNeurophysiol 90:1965–

583   1982.

584   Burton H, Sinclair RJ, Agato A (2012) Recognition memory for Braille or spoken words: an fMRI

585   study in early blind. Brain Res 1438:22–34

586   Burton H, Snyder AZ, Diamond JB, Raichle ME (2002) Adaptive changes in early and late blind:

587   a FMRI study of verb generation to heard nouns. JNeurophysiol 88:3359–3371

588   Calvert GA, Bullmore E, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PWR,

589   Iversen SD, David AS (1997) Activation of auditory cortex during silent lipreading. Science

590   (80-) 276:593–596.

591   Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., Ghazanfar, A. a., 2009. The natural

592   statistics of audiovisual speech. PLoS Comput. Biol. 5. doi:10.1371/journal.pcbi.1000436

593   Chomsky, N. (1975). Reflections on language. New York: Parthenon Press.

594   Collignon O, Dormal G, Albouy G, Vandewalle G, Voss P, Phillips C, Lepore F (2013) Impact of

595   blindness onset on the functional organization and the connectivity of the occipital cortex.

596   Brain 136:2769–2783

597   Collignon O, Dormal G, de Heering A, Lepore F, Lewis TL, Maurer D (2015) Long-Lasting

598   Crossmodal Cortical Reorganization Triggered by Brief Postnatal Visual Deprivation. Curr

599   Biol 25:2379–2383

600   Collignon O, Dormal G, Lepore F (2012) Building the brain in the dark : Functional and specific

601   crossmodal reorganization in the occipital cortex of blind individuals. In: Plasticity in Sensory

602   Systems (Jenkin M, Steeves JKE, Harris LR, eds), pp 73–96. Cambridge University Press.

31

603 Collignon O, Vandewalle G, Voss P, Albouy G, Charbonneau G, Lassonde M, Lepore F (2011)

604   Functional specialization for auditory-spatial processing in the occipital cortex of

605   congenitally blind humans. Proc Natl Acad Sci U S A 108:4435–4440

606 Collignon, O., Voss, P., Lassonde, M., Lepore, F., 2009. Cross-modal plasticity for the spatial

607   processing of sounds in visually deprived subjects. Exp. brain Res. 192, 343–58.

608   doi:10.1007/s00221-008-1553-z

609 Davis MH, Johnsrude IS (2003) Hierarchical processing in spoken language comprehension. J

610   Neurosci 23:3423–3431.

611 de Heer WA, Huth AG, Griffiths TL, Gallant JL, Theunissen FE (2017) The hierarchical cortical

612   organization of human speech processing. J Neurosci.

613 Deen B, Saxe R, Bedny M (2015) Occipital Cortex of Blind Individuals Is Functionally Coupled

614   with Executive Control Areas of Frontal Cortex. J Cogn Neurosci 27:1633–1647.

615 Dehaene-Lambertz G, Hertz-Pannier L, Dubois J (2006) Nature and nurture in language

616   acquisition: anatomical and functional brain-imaging studies in infants. Trends Neurosci

617   29:367–373.

618 Ding N, Simon JJZ (2014) Cortical entrainment to continuous speech: functional roles and

619   interpretations. Front Hum Neurosci 8:311.

620 Dormal G, Rezk M, Yakobov E, Lepore F, Collignon O (2016) Auditory motion in the sighted and

621   blind: Early visual deprivation triggers a large-scale imbalance between auditory and

622   "visual" brain regions. Neuroimage 134:630–644

623 Elman JL, Bates EA, Johnson MH, Karmiloff-Smith A, Parisi D, Plunkett K (1996) Rethinking

624   Innateness: Connectionist Perspective on Development.

625   Falchier A, Clavagnier S, Barone P, Kennedy H (2002) Anatomical evidence of multimodal

626   integration in primate striate cortex. JNeurosci 22:5749–5759.

627   Fonteneau E, Bozic M, Marslen-Wilson WD (2015) Brain network connectivity during language

628   comprehension: Interacting linguistic and perceptual subsystems. Cereb Cortex 25:3962–

629   3976.

630   Ghazanfar A, Schroeder C (2006) Is neocortex essentially multisensory? Trends Cogn Sci 10:278–

631   285.

632   Giordano BL, Ince RAA, Gross J, Schyns PG, Panzeri S, Kayser C (2017) Contributions of local

633   speech encoding and functional connectivity to audio-visual speech perception. Elife 6:1–

634   20.

635   Giraud A-L, Poeppel D (2012) Cortical oscillations and speech processing: emerging

636   computational principles and operations. Nat Neurosci 15:511–517

637   Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S (2013) Speech Rhythms

638   and Multiplexed Oscillatory Sensory Coding in the Human Brain. PLoS Biol 11.

639   Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R, Hämäläinen M (2001)

640   Dynamic Imaging of Coherent Sources: Studying Neural Interactions in the Human Brain.

641   Proc Natl Acad Sci 98:694–699.

642   Hannagan T, Amedi A, Cohen L, Dehaene-Lambertz G, Dehaene S (2015) Origins of the

643   specialization for letters and numbers in ventral occipitotemporal cortex. Trends Cogn Sci

644   19:374–382 Available at: http://dx.doi.org/10.1016/j.tics.2015.05.006.

645   Hasson U, Andric M, Atilgan H, Collignon O (2016) Congenital blindness is associated with large-

646   scale reorganization of anatomical networks. Neuroimage 128:362–372.

647   He C, Peelen M V, Han Z, Lin N, Caramazza A, Bi Y (2013) Selectivity for large nonmanipulable

648   objects in scene-selective visual cortex does not require visual experience. Neuroimage

649   79:1–9

650   Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci

651   8:393–402.

652   Horowitz-Kraus T, Grainger M, DiFrancesco M, Vannest J, Holland SK (2015) Right is not always

653   wrong: DTI and fMRI evidence for the reliance of reading comprehension on language-

654   comprehension networks in the right hemisphere. Brain Imaging Behav 9:19–31.

655   Jiang F, Stecker GC, Boynton GM, Fine I (2016) Early Blindness Results in Developmental

656   Plasticity for Auditory Motion Processing within Auditory and Occipital Cortex . Front Hum

657   Neurosci   10:324.

658   Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-Phase Coding Boosts and

659   Stabilizes Information Carried by Spatial and Temporal Spike Patterns. Neuron 61:597–608.

660   Kayser C, Petkov CI, Logothetis NK (2008) Visual Modulation of Neurons in Auditory Cortex.

661   Cereb Cortex 18:1560–1574.

662   Kayser SJ, Ince RAA, Gross J, Kayser C (2015) Irregular Speech Rate Dissociates Auditory Cortical

663   Entrainment, Evoked Responses, and Frontal Alpha. J Neurosci 35:14691–14701

664   Klinge C, Eippert F, Roder B, Buchel C (2010) Corticocortical Connections Mediate Primary Visual

665   Cortex Responses to Auditory Stimulation in the Blind. J Neurosci 30:12798–12805.

666   Kuhl PK, Meltzoff AN (1982) The bimodal perception of speech in infancy. Science (80-)

667   218:1138–1141.

668   Kushnerenko E, Teinonen T, Volein A, Csibra G (2008) Electrophysiological evidence of illusory

669   audiovisual speech percept in human infants. Proc Natl Acad Sci U S A 105:11442–11445.

670   Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain

671   signals. Hum Brain Mapp 8:194–208.

672   Lakatos P, Chen C, Connell MNO, Mills A, Schroeder CE (2007) Article Neuronal Oscillations and

673   Multisensory Interaction in Primary Auditory Cortex. :279–292.

674   Lamme VAF, Supèr H, Spekreijse H (1998) Feedforward, horizontal, and feedback processing in

675   the visual cortex. Curr Opin Neurobiol 8:529–535.

676   Lane C, Kanjlia S, Omaki A, Bedny M (2015) "Visual" Cortex of Congenitally Blind Adults

677   Responds to Syntactic Movement. J Neurosci 35:12859 LP-12868.

678   Lewkowicz DJ (2000) Infants' perception of the audible, visible, and bimodal attributes of

679   multimodal syllables. Child Dev 71:1241–1257.

680   Lewkowicz DJ (2010) Infant perception of audio-visual speech synchrony. Dev Psychol 46:66–77.

681   Lewkowicz DJ, Ghazanfar AA (2006) The decline of cross-species intersensory perception in

682   human infants. Proc Natl Acad Sci 103:6771–6774.

683   Lewkowicz DJ, Hansen-Tift AM (2012) Infants deploy selective attention to the mouth of a talking

684   face when learning speech. Proc Natl Acad Sci U S A 109:1431–1436.

685   Lewkowicz DJ, Leo I, Simion F (2010) Intersensory perception at birth: Newborns match

686   nonhuman primate faces and voices. Infancy 15:46–60.

687   Liljeström M, Kujala J, Jensen O, Salmelin R (2005) Neuromagnetic localization of rhythmic

688   activity in the human brain: a comparison of three methods. Neuroimage 25:734–745.

689   Luo H, Liu Z, Poeppel D (2010) Auditory Cortex Tracks Both Auditory and Visual Stimulus

690   Dynamics Using Low-Frequency Neuronal Phase Modulation. 8.

691   Luo H, Poeppel D (2007) Phase Patterns of Neuronal Responses Reliably Discriminate Speech in

692   Human Auditory Cortex. Neuron 54:1001–1010.

693   Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-

694   talker speech perception. Nature 485:233–236.

695   Murray MM, Lewkowicz DJ, Amedi A, Wallace MT (2016) Multisensory Processes: A Balancing

696   Act across the Lifespan. Trends Neurosci xx:1–13

697   Narain C, Scott SK, Wise RJS, Rosen S, Leff A, Iversen SD, Matthews PM (2003) Defining a Left-

698   lateralized Response Specific to Intelligible Speech Using fMRI. Cereb Cortex 13:1362–

699   1368.

700   Nath AR, Fava EE, Beauchamp MS (2011) Neural Correlates of Interindividual Differences in

701   Children's Audiovisual Speech Perception. J Neurosci 31:13963–13971.

702   Nolte G (2003) The magnetic lead field theorem in the quasi-static approximation and its use for

703   magnetoencephalography forward calculation in realistic volume conductors. Phys Med Biol

704   48:3637–3652.

705   Nolte G, Ziehe A, Nikulin V V., Schlögl A, Krämer N, Brismar T, Müller KR (2008) Robustly

706   estimating the flow direction of information in complex physical systems. Phys Rev Lett 100.

707   Obleser J, Eisner F, Kotz S a (2008) Bilateral speech comprehension reflects differential sensitivity

708   to spectral and temporal features. J Neurosci 28:8116–8123.

709   Okada K, Rong F, Venezia J, Matchin W, Hsieh IH, Saberi K, Serences JT, Hickok G (2010)

710   Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the

711   response to intelligible speech. Cereb Cortex 20:2486–2495.

712   Park H, Ince RAA, Schyns PG, Thut G, Gross J (2015) Frontal Top-Down Signals Increase Coupling

713   of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. Curr Biol

714   25:1649–1653.

715   Park H, Kayser C, Thut G, Gross J (2016) Lip movements entrain the observers' low-frequency

716   brain oscillations to facilitate speech intelligibility. Elife 5:e14521.

717   Patterson ML, Werker JF (2003) Two-month-old infants match phonetic information in lips and

718   voice. Dev Sci 6:191–196.

719   Peelen M, Bracci S, Lu X, He C, Caramazza A, Bi Y (2013) Tool Selectivity in Left Occipitotemporal

720   Cortex Develops without Vision. J Cogn Neurosci 25:1225–1234.

721   Peelle JE, Davis MH (2012) Neural oscillations carry speech rhythm through to comprehension.

722   Front Psychol 3:1–17.

723   Peelle JE, Gross J, Davis MH (2013) Phase-locked responses to speech in human auditory cortex

724   are enhanced during comprehension. Cereb Cortex 23:1378–1387.

725   Peelle JE, Johnsrude IS, Davis MH (2010) Hierarchical processing for speech in human auditory

726   cortex and beyond. Front Hum Neurosci 20:2486–2495.

727   Pellegrino F, Coupé C, Marisco E (2011) A Cross-Language Perspecitve on Speech Information

728   Rate. Language (Baltim) 87:539–558.

729   Pietrini P, Furey ML, Ricciardi E, Gobbini MI, Wu W-HCH, Cohen L, Guazzelli M, Haxby J V (2004)

730   Beyond sensory images: Object-based representation in the human ventral pathway. Proc

731   Natl Acad Sci U S A 101:5658–5663

732   Poeppel D (2003) The analysis of speech in different temporal integration windows: Cerebral

733   lateralization as "asymmetric sampling in time." In: Speech Communication, pp 245–255.

734   Poeppel D, Idsardi WJ, van Wassenhove V (2008) Speech perception at the interface of

735   neurobiology and linguistics. Philos Trans R Soc Lond B Biol Sci 363:1071–1086.

736   Poirier C, Collignon O, Scheiber C, Renier L, Vanlierde a, Tranduy D, Veraart C, De Volder a G

737   (2006) Auditory motion perception activates visual motion areas in early blind subjects.

738   Neuroimage 31:279–285 Available at: http://www.ncbi.nlm.nih.gov/pubmed/16443376

739   [Accessed November 6, 2014].

740   Pons F, Lewkowicz DJ, Soto-Faraco S, Sebastián-Gallés N (2009) Narrowing of intersensory

741   speech perception in infancy. Proc Natl Acad Sci U S A 106:10598–10602.

742   Reich L, Szwed M, Cohen L, Amedi A (2011) A ventral visual stream reading center independent

743   of visual experience. Curr Biol 21:363–368

744   Ricciardi E, Vanello N, Sani L, Gentili C, Scilingo EP, Landini L, Guazzelli M, Bicchi A, Haxby J V,

745   Pietrini P (2007a) The effect of visual experience on the development of functional

746   architecture in hMT+. CerebCortex 17:2933–2939

747   Ricciardi E, Vanello N, Sani L, Gentili C, Scilingo EP, Landini L, Guazzelli M, Bicchi A, Haxby J V,

748   Pietrini P (2007b) The effect of visual experience on the development of functional

749   architecture in hMT+. Cereb Cortex 17:2933–2939.

750   Rodd JM, Davis MH, Johnsrude IS (2005) The neural mechanisms of speech comprehension:

751   fMRI studies of semantic ambiguity. Cereb Cortex 15:1261–1269.

752   Rodd JM, Johnsrude IS, Davis MH (2010) The role of domain-general frontal systems in language

753   comprehension: Evidence from dual-task interference and semantic ambiguity. Brain Lang

754   115:182–188.

755   Roder B, Stock O, Bien S, Neville H, Rosler F (2002) Speech processing activates visual cortex in

756   congenitally blind humans. EurJNeurosci 16:930–936

757   Rosenblum LD, Schmuckler MA, Johnson JA (1997) The McGurk effect in infants. Percept

758   Psychophys 59:347–357.

759   Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2007) Do you see what I am saying?

760   Exploring visual enhancement of speech comprehension in noisy environments. Cereb

761   Cortex 17:1147–1153.

762   Sadato N, Pascual-Leone A, Grafman J (1996) activation of the primary visual cortex by braille

763   reading in blind subjects. Nature.

764   Schepers IM, Hipp JF, Schneider TR, Röder B, Engel AK (2012) Functionally specific oscillatory

765   activity correlates between visual and auditory cortex in the blind. Brain 135:922–934.

766   Schroeder C, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory

767   selection. Trends Neurosci 32:1–16.

768   Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A (2008) Neuronal oscillations and visual

769   amplification of speech. Trends Cogn Sci 12:106–113.

770   Schwartz JL, Savariaux C (2014) No, There Is No 150 ms Lead of Visual Speech on Auditory

771   Speech, but a Range of Audiovisual Asynchronies Varying from Small Audio Lead to Large

772   Audio Lag. PLoS Comput Biol 10.

773   Scott SK, Blank CC, Rosen S, Wise RJS (2000) Identification of a pathway for intelligible speech

774   in the left temporal lobe. Brain 123:2400–2406.

775   Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of

776   smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44:83–

777   98.

778   Striem-amit E, Ovadia-caro S, Caramazza A, Margulies S, Villringer A, Amedi A (2015) Functional

779   connectivity of visual cortex in the blind follows retinotopic organization principles. :1–17.

780 Szymanski FD, Rabinowitz NC, Magri C, Panzeri S, Schnupp JWH (2011) The Laminar and

781   Temporal Structure of Stimulus Information in the Phase of Field Potentials of Auditory

782   Cortex. J Neurosci 31:15787–15801.

783 Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A (1997) Localization of brain electrical

784   activity via linearly constrained minimum variance spatial filtering. IEEE Trans Biomed Eng

785   44:867–880.

786 van Wassenhove V (2013) Speech through ears and eyes: Interfacing the senses with the

787   supramodal brain. Front Psychol 4.

788 Vouloumanos A, Druhen MJ, Hauser MD, Huizink AT (2009) Five-month-old infants' identification

789   of the sources of vocalizations. Proc Natl Acad Sci U S A 106:18867–18872.

790 Weeks R, Horwitz B, Aziz-Sultan a, Tian B, Wessinger CM, Cohen LG, Hallett M, Rauschecker JP

791   (2000) A positron emission tomographic study of auditory localization in the congenitally

792   blind. J Neurosci 20:2664–2672.

793 Weikum WM, Vouloumanos A, Navarra J, Soto-Faraco S, Sebastián-Gallés N, Werker JF (2007)

794   Visual language discrimination in infancy. Science 316:1159.

795 Zion Golumbic E, Cogan GB, Schroeder CE, Poeppel D (2013a) Visual input enhances selective

796   speech envelope tracking in auditory cortex at a "cocktail party". J Neurosci 33:1417–1426

797 Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR,

798   Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013b) Mechanisms underlying

799   selective neuronal tracking of attended speech at a "cocktail party." Neuron 77:980–991

800 Zoefel B, VanRullen R (2015) Selective Perceptual Phase Entrainment to Speech Rhythm in the

801   Absence of Spectral Energy Fluctuations. J Neurosci 35:1954–1964.

802

803

804

Author Contributions

806   MJVA, OC, and FB designed the experiment. MJVA, FB, RB, and SM screened and tested the

807   participants. MJVA analysed the data. MJVA and OC wrote the manuscript. FB, RB and SM did

808   comments on the draft of the manuscript.

809   **Figures**
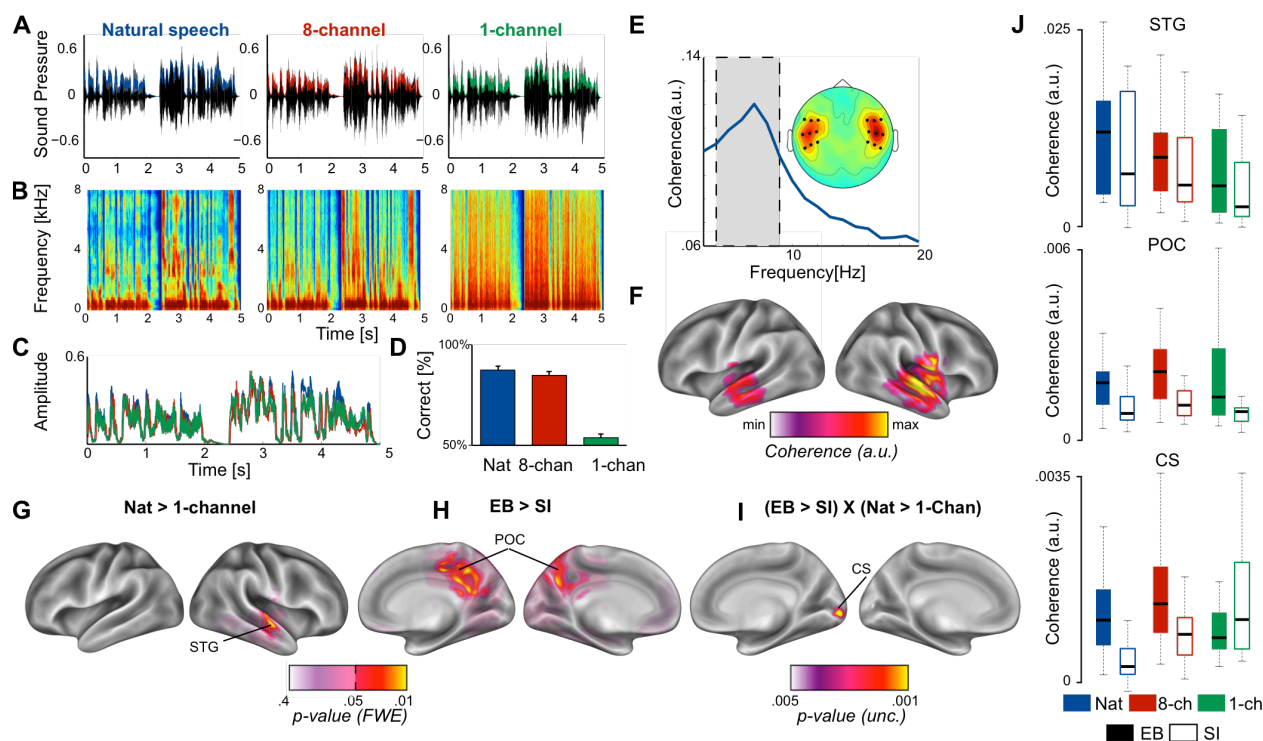
810   **Figure 1**

811
812

813   *Figure 1.* Cerebro-acoustic coherence in EB and SI. (A) Segments from each condition (nat, 8-

814   channel, and 1-channel) in the time domain. Overlaid is the amplitude envelope of each

815   condition. (B) Spectrogram of the speech samples from (A) The effect of spectral vocoding on

816   the fine spectral detail in all three conditions. (C) The amplitude envelopes of the samples for

817   each condition superimposed are highly similar, despite distortions in the fine spectral detail. (D)

818   Behavioural performance on the comprehension statements reveals that comprehension is

819   unperturbed in the nat and 8-channel conditions, whereas the 1-channel conditions elicits chance

820   performance. (E) Coherence spectrogram extracted from bilateral temporal sensors reveals a

821   peak in cerebro-acoustic coherence at 7Hz, across groups and conditions. The shaded area

822   depicts the frequency range of interest in the current study. The topography shows the spatial

823   extent of coherence values in the range 4-9Hz. Enhanced coherence is observed selectively in

824   bilateral temporal sensors. (F) Source reconstruction of the raw coherence confirms that cerebro-

825   acoustic coherence is strongest in the vicinity of auditory cortex, bilaterally, extending into

826   superior and middle temporal gyrus. (G) Statistically thresholded maps for the contrast between

827   natural and 1-channel vocoded speech show an effect of intelligibility (p<.05, FWE-corrected) in

828   right STG. (H) Enhanced envelope tracking is observed for EB versus SI in a bilateral parieto-

829   occipital network network along the medial wall centred on Precuneus (p<.05, FWE-corrected).

830   (I) The statistical map shows the interaction effect between blindness and intelligibility: Early

831   blind individuals show enhanced synchronization during intelligible (nat) versus non-intelligible

832   speech (1-channel) as compared to SI in right calcarine sulcus (CS) (p<.005, uncorrected). (J)

833   Boxplots for three regions identified in the whole brain analysis (top panel: STG; middle panel:

834   parieto-occipital cortex; bottom panel: calcarine sulcus).
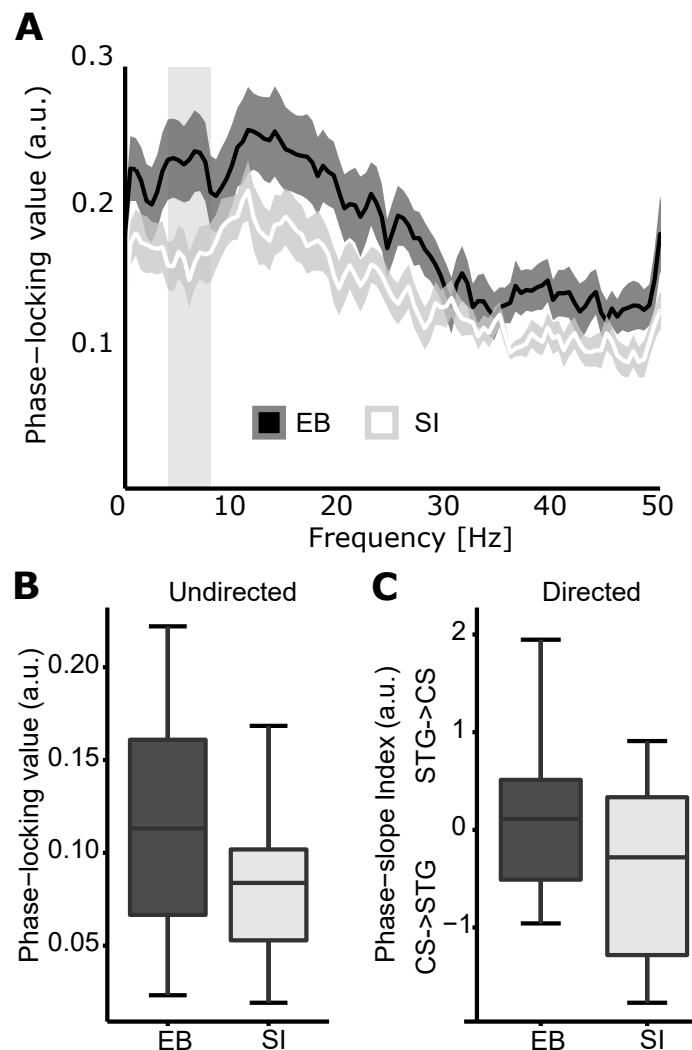
835

836   Figure 2

837

838   *Figure 2.* Occipital-temporal connectivity in EB and SI. (A) Spectra illustrate the phase locking

839   between CS and STG in EB (dark curve), and SI (light curve). Shaded areas illustrate the SEM. A

840   difference between the curves is observed in the theta range. (B) Boxplots depict the mean phase

841   locking between CS and STG for EB (black), and SI (white) in the theta range. Connectivity is

842   enhanced in EB versus SI. (C) Directional connectivity analysis using the phase-slope index (PSI).

843   Positive values suggest enhanced directionality from STG to CS, and negative values represent

844   enhanced connectivity in the opposite direction. He boxplots highlight that SI show a strong

845   feed-forward drive from CS to STG, whereas blind individuals show a more balanced pattern,

846   with a non-significant trend in the opposite direction.

847
848
849
850

851   **Supplementary materials**

852   S1. Movie of the natural and two vocoding conditions