PRIFYSGOL
BANGOR
UNIVERSITY

High-level language processing regions are not engaged in action observation or imitation

Pritchett, Brianna L; Hoeflin, Caitlyn; Koldewyn, Kami; Dechter, Eyal; Fedorenko, Evelina

Journal of Neurophysiology

Cyswllt i'r cyhoeddiad / Link to publication

1 **Title**
2 High-level language processing regions are not engaged in action observation or
3 imitation
4
5 **Abbreviated title**
6 Action response in language regions
7
8 **Authors**
9 Brianna L. Pritchett[1], Caitlyn Hoeflin[1], Kami Koldewyn[2], Eyal Dechter[1], and Evelina
10 Fedorenko[1,3,4]
11
12 **Affiliations**
13 [1] Department of Brain and Cognitive Sciences / McGovern Institute for Brain Research,
14 Massachusetts Institute of Technology, Cambridge, MA 02139, USA
15 [2] School of Psychology, Bangor University, Gwynedd LL57 2AS, U.K.
16 [3] Department of Psychiatry, Massachusetts General Hospital, Charlestown, MA 02129,
17 USA
18 [4] Department of Psychiatry, Harvard Medical School, Boston, MA 02115, USA
19
20 **Corresponding Author**
21 Evelina Fedorenko (evelina9@mit.edu; 43 Vassar Street, 46-3037F, Cambridge MA
22 02139)
23

30
31 **Abstract (234 words)**
32
33 A set of left frontal, temporal, and parietal brain regions respond robustly during

34 language comprehension and production (e.g., Fedorenko et al. 2010; Menenti et al.

35 2011). These regions have been further shown to be selective for language relative to

36 other cognitive processes, including arithmetic, aspects of executive function, and music

37 perception (e.g., Fedorenko et al. 2011; Monti et al. 2012). However, one claim about

38 overlap between language and non-linguistic cognition remains prominent. In particular,

39 some have argued that language processing shares computational demands with action

40 observation and/or execution (e.g., Rizzolatti and Arbib 1998; Koechlin and Jubault

41 2006; Tettamanti and Weniger 2006). However, the evidence for these claims is indirect,

42 based on observing activation for language and action tasks within the same broad

43 anatomical areas (e.g., on the lateral surface of the left frontal lobe). To test whether

44 language indeed shares machinery with action observation/execution, we examined the

45 responses of language brain regions, defined functionally in each individual participant

46 (Fedorenko et al. 2010), to action observation (Experiments 1, 2, 3a) and action imitation

47 (Experiment 3b). With the exception of the language region in the angular gyrus, all

48 language regions, including those in the inferior frontal gyrus (within "Broca's area"),

49 showed little or no response during action observation/imitation. These results add to the

50 growing body of literature suggesting that high-level language regions are highly

51 selective for language processing (see Fedorenko and Varley 2016 for a review).

52

53

2

54   **New & Noteworthy** (75 words)

55   Many have argued for overlap in the machinery used to interpret language and others'

56   actions, either because action observation was a precursor to linguistic communication or

57   because both require interpreting hierarchically-structured stimuli. However, existing

58   evidence is indirect, relying on group analyses or reverse inference. We examined

59   responses to action observation in language regions defined functionally in individual

60   participants and found no response. Thus, language comprehension and action

61   observation recruit distinct circuits in the modern brain.

62

63 **Introduction**

64 Although brain regions that support high-level language processing have been shown to

65 be selective for language over various non-linguistic cognitive processes (e.g., Fedorenko

66 and Varley 2016), the idea of overlap between language processing and action

67 observation and/or execution remains prominent in the literature. Two lines of theorizing

68 have been used to argue for this overlap. The first stemmed from the discovery of mirror

69 neurons in the prefrontal cortex of rhesus macaques. These neurons fire both when a

70 monkey performs an action and when it observes the action performed (Rizzolatti et al.

71 1988). Rizzolatti & Arbib (1998; Arbib 2005, 2010; see also Petrides & Pandya 2009;

72 Corballis 2010) speculated that in our primate ancestors, mirror neurons were critical for

73 understanding one another's actions – a core component of social cognition. They argued

74 that, over time, basic manual actions grew more abstract, and eventually became signs,

75 which, in turn, became mouth movements/vocalizations. Thus, manual actions are argued

76 to be a fundamental precursor to linguistic communication, and action understanding and

77 language comprehension should share a common neural substrate because they share a

78 common functional ancestor.

79   Although the general idea that language arose from gesture finds substantial

80 support (e.g., Tomasello 2008; Corballis 2003; cf. Slocombe 2015), the role of mirror

81 neurons in the evolution of language remains debated (e.g., Hickok 2009). The existence

82 of brain cells / regions with properties of the macaque mirror neuron system in humans is

83 supported by a number of studies (e.g., Mukamel et al. 2010; see Molenberghs et al. 2012

84 for a meta-analysis) but has not gone unchallenged (e.g., Dinstein et al. 2007; Lingnau et

85 al. 2009). Regardless of these controversies, however, given the prominence of the

4

86 gesture-based hypothesis of language evolution, it seems important to test whether any

87 parts of the language network in the modern human brain respond to action

88 observation/execution.

89   The second general line of reasoning is that both the language system and the

90 action observation system (possibly restricted to biological actions; e.g., Clerget et al.

91 2009; Fazio et al. 2009) rely on an amodal mechanism that recognizes and produces

92 hierarchical structure (e.g., Fiebach and Schuboltz 2006; Koechlin and Jubault 2006;

93 Tettamanti and Weniger 2006). This mechanism has been argued to reside in the inferior

94 frontal gyrus (IFG), in or around "Broca's area" (we use quotations because the definition

95 of this brain region in the literature is extremely variable, and the term has been argued

96 by some to no longer be meaningful as a result; Tremblay & Dick, 2016). However, the

97 evidence for overlap between language and action observation in the IFG is problematic

98 because the IFG is among the most structurally (e.g., Amunts et al. 2010) and

99 functionally (e.g., Fedorenko et al. 2012a) heterogeneous brain regions. Further, lateral

100 frontal lobes are characterized by high inter-individual variability (e.g., Amunts et al.

101 1999; Tomaiuolo et al. 1999; Juch et al. 2005). Thus, activation overlap between

102 language and action observation in a traditional fMRI group analysis (e.g., Higuchi et al.

103 2009), where activations are averaged across individuals, can be misleading (e.g., Nieto-

104 Castañon and Fedorenko 2012), particularly in the aforementioned regions.

105   Further, some prior studies did not even include a direct within-experiment

106 comparison between a language and an action task (e.g., Binkofsky et al. 2000; Meister

107 and Iacaboni 2007; Clerget et al. 2009) and relied solely on the fallacious reverse

108 inference (Poldrack 2006, 2011) to interpret the frontal activations for action tasks. This

5

109   approach is especially problematic in this case because frontal lobes, including "Broca's

110   area" itself (Fedorenko et al. 2012a), contain both i) language-selective regions, and ii)

111   highly domain-general ones that belong to the fronto-parietal multiple demand (MD)

112   network (e.g., Duncan 2010) and are driven by diverse cognitive demands (e.g., Duncan

113   & Owen 2000; Fedorenko et al. 2013). Thus, interpreting frontal activations for an action

114   observation task as reflecting the recruitment of the language system is not justified.

115   Similarly, although many aphasic patients with frontal lesions exhibit deficits in action

116   observation/execution (e.g., Kimura 1977; Kimura et al. 1976; Papagno et al., 1993;

117   Saygin et al. 2004), these patients' lesions are often extensive and plausibly affect two or

118   more functionally distinct regions (cf. Sirugu et al. 1998). Thus, arguing for overlap in

119   mechanisms that support language processing and action observation based on such data

120   is also not warranted.

121   To test – in the most direct way – whether action observation/execution relies on

122   some of the same neural mechanisms as high-level language processing, we examined

123   responses to action observation and imitation in the language regions functionally defined

124   in each individual. This analytic approach circumvents the problem of high inter-

125   individual variability in the precise locations of functional regions (e.g., Fischl et al.

126   2008; Frost and Goebel 2011; Tahmasebi et al. 2011) and thus stands a chance to

127   conclusively answer the question about whether language regions support some aspects

128   of action observation. It is worth noting that this question is conceptually distinct from

129   the question that is at the core of the embodiment debate (see Leshinskaya & Caramazza

130   2016 for a recent review): namely, whether concepts are "grounded" in sensory-motor

131   systems. We elaborate further on the relationship between these questions in the

6

132   Discussion.

133

134   **Materials and Methods**

135   The general approach adopted here across the four experiments is as follows: first, we

136   identify the language network in each participant individually using a functional localizer

137   task based on a broad contrast between the reading of sentences vs. sequences of

138   nonwords (Fedorenko et al. 2010). Then, we examine the engagement of these language-

139   responsive voxels in action observation/imitation across several paradigms. This

140   approach has been previously shown to yield higher sensitivity and functional resolution

141   than traditional group-based analyses, as well as more accurate estimates of effect sizes

142   (e.g., Saxe et al. 2006; Nieto-Castañon and Fedorenko 2012). Further, this approach

143   makes the results directly comparable across the four experiments.

144   It is worth emphasizing that we here focus on *high-level* language processing

145   regions, i.e., brain regions that support lexico-semantic and combinatorial (semantic and

146   syntactic) processing (e.g., Fedorenko et al. 2012b; Bautista and Wilson 2016; Blank et al.

147   2016). These regions plausibly underlie our ability to infer meanings from others'

148   linguistic utterances during comprehension as well as to convert our thoughts into

149   linguistic forms during production. This high-level language network is distinct from

150   both lower-level *perceptual* regions that respond selectively to speech, but are not

151   sensitive to the meaningfulness of the speech signal (e.g., Overath et al. 2015; Norman-

152   Haignere et al. 2015) and lower-level *speech articulation* regions that respond robustly

153   when we produce speech sounds, but again are not sensitive to the meaningfulness of the

154   utterance (e.g., Bohland and Guenther, 2006; Flinker et al. 2015; Basilakos et al. 2017).

7

155 Thus, our main conclusions pertain to the high-level component of the extended language

156 network. We return to this issue in the Results section.

157

158 **Participants.** Participants were recruited from MIT and the surrounding

159 Cambridge/Boston, MA community and were paid for their participation. Eleven

160 participants were tested in Experiment 1, 57 in Experiment 2, 13 in Experiment 3a, and

161 16 in Experiment 3b. Seven participants were excluded (3 for excessive motion – all in

162 Experiment 3b, 2 for equipment failure, 1 because an incorrect scanner sequence was

163 used, and 1 due to experimenter error), leaving 90 participants for analysis (10 in

164 Experiment 1, 54 in Experiment 2, 13 in Experiment 3a, and 13 in Experiment 3b). (The

165 number of participants in Experiment 2 was so large because this experiment was used

166 across multiple projects, and we decided to include here all the data available.) Due to

167 some overlap in participants across experiments (8 participated in both Experiment 2 and

168 3a, and 5 participated in both Experiment 2 and 3b), there were 77 unique individuals

169 (age 18-52, mean age 24, 43 females), 68 right-handed (as determined by the Edinburgh

170 handedness inventory, Oldfield 1971, for n=69, or self report). No participants were

171 excluded based on handedness because we would like to generalize our results to the

172 entire population, as opposed to only the right-handed participants (see Willems et al.

173 2014, for discussion). The nine left-handed participants all had a left-lateralized language

174 network, as determined by the language localizer task described below. To determine

175 lateralization, the number of language-contrast-activated voxels in the right hemisphere at

176 a fixed significance threshold was subtracted from the number of language voxels in the

177 left hemisphere at the same threshold, and the resulting value was divided by the sum of

8

178   language voxels across hemispheres (see Mahowald & Fedorenko 2016 for further details

179   on this method). All were native speakers of English, had normal hearing and vision, and

180   no history of language impairment. The protocol for these studies was submitted to, and

181   approved by, MIT's Committee on the Use of Humans as Experimental Subjects

182   (COUHES). All participants gave written informed consent in accordance with the

183   requirements of this protocol.

184

185   **Design and procedure common to all four experiments.** Each participant completed a

186   language localizer task (Fedorenko et al. 2010) and an action observation/imitation task.

187   12 participants completed the localizer task in a separate scanning session; the remaining

188   78 participants performed the localizer and an action experiment in the same session,

189   along with one or two additional tasks for unrelated studies. The entire scanning session

190   lasted for approximately 2 hours. The task used to localize the language network is

191   described in detail in Fedorenko et al. (2010); the materials and scripts are available from

192   the Fedorenko Lab website (https://evlab.mit.edu/funcloc). Briefly, we used a reading

193   task contrasting sentences (e.g., THE SPEECH THAT THE POLITICIAN PREPARED

194   WAS TOO LONG FOR THE MEETING) and lists of unconnected, pronounceable

195   nonwords (e.g., LAS TUPING CUSARISTS FICK PRELL PRONT CRE POME

196   VILLPA OLP WORNETIST CHO) in a standard blocked design with a counterbalanced

197   order across runs (for timing parameters, see Table 1). The sentences > nonwords

198   contrast targets brain regions that support lexico-semantic and combinatorial (semantic

199   and syntactic) processing. Stimuli were presented one word/nonword at a time. For 10

200   participants (in Experiment 1), each trial ended with a memory probe and they had to

9

201   indicate, via a button press, whether or not that probe had appeared in the preceding

202   sequence of words/nonwords. The remaining participants instead read the materials

203   passively (we included a button-pressing task at the end of each trial, to help participants

204   remain alert). Importantly, this localizer has been shown to generalize across task

205   manipulations: the sentences > nonwords contrast, and similar contrasts between

206   language and a linguistically degraded control condition, robustly activates the fronto-

207   temporal language network regardless of the task, materials, and modality of presentation

208   (e.g., Fedorenko et al. 2010; Fedorenko 2014; Scott et al. 2016).

209   The action observation tasks included a variety of conditions – including hand

210   actions with (Experiment 1) or without (Experiment 3a) a manipulable object, actions

211   that involve different body parts including hands, but also arms, legs, feet, torso, and

212   head (Experiment 2), face actions (Experiments 2 and 3a), and specifically eye and

213   mouth actions (Experiment 3a); the action imitation task similarly included several

214   conditions (Experiment 3b). We describe each experiment in more detail below.

215

216   ***Experiment 1: Hand action observation***

217   Participants watched short videos where a small non-nameable 3D object was

218   manipulated in some way by a hand, in a blocked design, and performed a simple one-

219   back task designed to draw attention to the action or the object. (We used non-nameable

220   objects to avoid a potential confound of activating the names of common objects, which

221   would likely elicit some response in the language regions, but not due to overlap in

222   computational demands between language understanding and action observation.) In the

223   action condition, participants had to press a button when they saw the same action twice

10

224   in a row, and in the object condition, they watched the same videos but had to press a

225   button when they saw the same object twice in a row. The task manipulation was

226   included in an effort to maximally focus the participants' attention on the actions in the

227   action condition.

228   *Materials*

229   There were 8 possible hand actions (e.g., push forward with back of the fingers, or pick

230   up with an index finger and a thumb) and 8 possible non-nameable objects, resulting in

231   64 unique stimuli (see Figure 1 for screenshots from sample stimuli). A short video was

232   created for each action/object combination. Each video started with the object sitting on a

233   table, and then the hand entered the frame (always from the same side), performed the

234   action, and exited the frame. Because objects take less time to identify than actions (given

235   that actions unfold over time), some steps were taken to make the conditions comparable

236   in difficulty. First, the videos were edited so that the action started as quickly as possible

237   after the onset of the video (on average, the action took about 250 ms to initiate). Second,

238   objects were grouped into "families" for presentation purposes such that objects within a

239   family were visually similar to one another. Conversely, actions were grouped in a way

240   such that actions within a set were visually dissimilar.

241   *Procedure*

242   Each video (trial) lasted 3 seconds, and trials were grouped into blocks of 8 trials each.

243   Each block was preceded by a 2-second instructions screen telling participants which

244   condition they were about to see. Each run consisted of 16 such experimental blocks (26

245   seconds each; 8 blocks per condition) and 5 fixation blocks (16 seconds each, placed at

246   the beginning of the run, and after each set of four blocks). Each run thus lasted 496

11

247   seconds (8 min 16 sec). Each participant saw either 4 or 5 runs. The order of conditions

248   was counterbalanced across runs and participants.

249

250   ***Experiment 2: Face and body action observation***

251   Participants passively watched silent videos of i) face actions, ii) body actions, iii)

252   driving through natural scenes, iv) moving man-made objects, and v) spatially scrambled

253   versions of these objects in a blocked design (see Pitcher et al., 2011, for a detailed

254   description). For the purposes of the current study, we examined the first two conditions:

255   face actions and body actions. Participants were instructed to watch attentively.

256   *Materials*

257   There were 60 unique stimuli per condition. The videos depicted children moving against

258   a black background. These children performed a variety of actions like dancing, walking,

259   and crawling (see Figure 1 for screenshots from sample stimuli). The face action videos

260   featured a child's face in motion – smiling, laughing, talking, or looking at someone off-

261   camera. The body action videos featured a child's moving body part – hands, arms, legs,

262   feet, torso, or back of the head – but did not include the face.

263   *Procedure*

264   Each trial consisted of a single video that lasted 3 seconds, and trials were grouped into

265   blocks of 6 trials each. Each run consisted of 10 experimental blocks (18 seconds each; 2

266   blocks per condition) and 3 fixation blocks (18 seconds each), placed at the beginning,

267   middle, and end of the run. Each run thus lasted 234 seconds (3 min 54 sec). Each

268   participant saw between 2 and 4 runs.

269

12

270 ***Experiment 3 a/b: Face, eye, mouth, and hand action observation / imitation***

271 Participants watched silent videos of an actress performing face actions, eye actions,

272 mouth actions, and hand actions. Additionally, the experiment included videos where the

273 actress pronounced consonant and vowel sounds, syllables, nonwords, and words in

274 English and German, and sang or hummed nonwords, all in a blocked design. For the

275 purposes of the current study, we examined the first four conditions: face actions, eye

276 actions, mouth actions, and hand actions. In the observation version of the experiment

277 (Experiment 3a), participants were asked to just watch attentively, and in the imitation

278 version (Experiment 3b), a different set of participants were instructed to imitate each

279 action while keeping their head as still as possible.

280 *Materials*

281 There were 8 unique stimuli per condition. The videos depicted a female actress against a

282 grey background. In the face, eye, and mouth action conditions, she was sitting facing the

283 camera, with the frame going from just below her shoulders to just above the top of her

284 head. Each video started and ended with the actress looking at the camera, with a neutral

285 expression. The face condition included actions like looking surprised or making a "fish"

286 face (see Figure 1 for screenshots from sample stimuli); the eye condition included

287 actions like moving the eyes up or to the lower left; and the mouth condition included

288 actions like touching the upper teeth with the tongue or pursing the lips to blow air out. In

289 the hand action condition, the hand rested on a wooden table, with the frame covering the

290 hand and a portion of the forearm. Each video started and ended with the hand resting on

291 the table. The hand condition included actions like pulling in the fingers or tapping a

292 finger or multiple fingers on the table.

293   *Procedure – Experiment 3a (observation)*

294   Each video (trial) lasted 5 seconds, and trials were grouped into blocks of 3 trials each.

295   Each run consisted of 26 experimental blocks (15 seconds each; 2 blocks for each of

296   thirteen conditions) and 5 fixation blocks (14 seconds each), placed at the beginning and

297   end of each run, as well as after the 7th, 13th, and 20th blocks. Each run thus lasted 460

298   seconds (7 min 30 sec). Each participant saw between 4 and 6 runs.

299   *Procedure – Experiment 3b (imitation)*

300   The procedure was identical to that of Experiment 3a except that each video (trial) lasted

301   8 seconds (5 seconds for the video and 3 seconds for the participant to imitate the action;

302   note that although the videos lasted 5 seconds each, the actual action does not take longer

303   than ~3 seconds). Each run thus lasted 694 seconds (11 min 34 sec). Each participant saw

304   between 3 and 8 runs.

305

306   **fMRI data acquisition and preprocessing.** Structural and functional data were collected

307   on the whole-body 3 Tesla Siemens Trio scanner with a 32-channel head coil at the

308   Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at

309   MIT. T1-weighted structural images were collected in 128 axial slices with 1 mm

310   isotropic voxels (TR = 2530 ms, TE = 3.48 ms). Functional, blood oxygenation level

311   dependent (BOLD) data were acquired using an EPI sequence (with a 90 degree flip

312   angle and using GRAPPA with an acceleration factor of 2), with the following

313   acquisition parameters: thirty-one 4 mm thick near-axial slices, acquired in an interleaved

314   order with a 10% distance factor, 2.1 mm x 2.1 mm in-plane resolution; field of view of

315   200 ms in the phase encoding anterior to posterior (A > P) direction; matrix size of 96

14

316   mm x 96 mm; TR of 2000 ms; and TE of 30 ms. Prospective acquisition correction

317   (Thesen et al. 2000) was used to adjust the positions of the gradients based on the

318   participant's motion from the previous TR. The first 10s of each run (before the start of

319   presentation of the stimuli) were excluded to allow for steady-state magnetization.

320   MRI data were analyzed using SPM5 and custom MATLAB and shell scripts.

321   Each participant's data were motion corrected, normalized into a common brain space

322   (MNI) and resampled into 2 mm isotropic voxels. The data were smoothed with a 4mm

323   Gaussian filter and high-pass filtered (at 200s). All task effects were estimated using a

324   General Linear Model (GLM) in which each experimental condition was modeled with a

325   boxcar function convolved with the canonical hemodynamic response function (HRF).

326

327   **Definition of group-constrained, subject-specific fROIs.** The critical analyses were

328   restricted to individually defined language fROIs (functional regions of interest). These

329   fROIs were defined using the Group-constrained Subject-Specific (GSS) approach

330   (Fedorenko et al. 2010; Julian et al. 2012) where a set of spatial parcels (binary masks

331   that correspond to locations where activation has been previously observed for the

332   relevant localizer contrast) is combined with each individual subject's localizer activation

333   map, to constrain the definition of individual fROIs. The parcels are sufficiently large to

334   encompass the extent of variability in the locations of individual activations. For the

335   critical language fROIs, we used a set of six parcels derived from a group-level

336   probabilistic activation overlap map for the sentences > nonwords contrast in 220

337   participants. These parcels (Figure 2) included three regions in the left frontal cortex: two

338   in the left inferior frontal gyrus (LIFG, LIFGorb), and one in the left middle frontal gyrus

15

339 (LMFG), two in the left temporal lobe (LAntTemp and LPostTemp), and one extending

340 into the angular gyrus (LAngG). These parcels are similar to the ones originally reported

341 in Fedorenko et al. (2010) based on a probabilistic activation overlap map from 25

342 participants, except that the two anterior temporal parcels were grouped together (the

343 original LAntTemp merged with LMidAntTemp), and the two posterior temporal parcels

344 were grouped together (the original LMidPostTemp merged with LPostTemp). The

345 parcels are available for download from https://evlab.mit.edu/funcloc.

346   Within each parcel, we selected the top 10% most responsive voxels, based on the

347 *t* values for the sentences > nonwords contrast (see e.g., Figure 1 in Blank et al. 2014; or

348 Figure 1 in Mahowald and Fedorenko 2016, for sample fROIs). Statistical tests were

349 performed on these values.

350   In addition to the language fROIs, a set of control fROIs was defined in the

351 participants in Experiments 2, 3a, and 3b. In particular, we used 18 anatomical parcels

352 across the two hemispheres (Tzourio-Mazoyer et al. 2002) covering frontal and parietal

353 brain areas that belong to the so-called multiple demand (MD) network (Duncan 2010,

354 2013). This network has been linked to executive demands across domains (e.g., Duncan

355 and Owen 2000; Fedorenko et al. 2013; Hugdahl et al. 2015), but parts of this network

356 have also been implicated in the processing of actions (e.g., Culham and Valear 2006;

357 Gallivan and Culham 2015; Biagi et al. 2015; Caspers et al. 2010). We thus expected

358 some of these regions to respond to action observation and/or imitation. In particular, we

359 focused on a subset of 6 parcels (although the results were corrected for the total number

360 of regions that comprise this network, i.e., 18): the bilateral IFGop and PrecG fROIs in

361 the frontal cortex because those lie in close proximity to the language fROIs, and the

16

362   bilateral SupPar fROIs in the parietal cortex because these regions have been implicated

363   in action observation in prior work (e.g., Johnson-Frey et al., 2005).

364   To define individual MD fROIs, we used a spatial working memory task where

365   participants keep track of locations within a grid (this MD localizer task was not included

366   in Experiment 1, hence this analysis could not be performed for those participants). The

367   task is described in detail in Fedorenko et al. (2013; see also Blank et al. 2014). Briefly,

368   on each trial, participants saw a 3×4 grid and kept track of eight (hard version) or four

369   (easy version) locations that were sequentially flashed two at a time or one at a time,

370   respectively. Then, participants indicated their memory for these locations in a two-

371   alternative, forced-choice paradigm via a button press. Feedback was provided after

372   every trial. Hard and easy conditions were presented in a standard blocked design (4 trials

373   in a 32s block, 6 blocks per condition per run) with a counterbalanced order across runs.

374   Each run included 4 blocks of fixation (16s each) and lasted a total of 448s. Within each

375   anatomical parcel, we selected the top 10% most responsive voxels, based on the *t* values

376   for the hard > easy spatial working memory contrast. Statistical tests were performed on

377   these values.

378   Finally, for some additional analyses reported in the Discussion, we examined i)

379   brain regions in the auditory cortex that support speech perception, and ii) brain regions

380   in the premotor cortex that support speech articulation. For the former, we used the

381   following anatomical parcels from the FSL atlas (Desikan et al. 2006): bilateral planum

382   polare (PP), planum temporale (PT), anterior superior temporal gyrus (ASTG), and

383   posterior superior temporal gyrus (PSTG). To define individual speech-responsive fROIs,

384   these anatomical parcels were masked with activation maps for a contrast between

17

385   listening to nonwords and observing hand actions (in Experiment 3a). The responses

386   were then extracted to nonwords, and the four action observation conditions. To estimate

387   the responses to the nonwords and hand action observation conditions, an across-runs

388   cross-validation procedure was used so that the data to define the fROIs and estimate

389   their responses were independent (e.g., Kriegeskorte et al. 2011). In particular, all but one

390   run were used to define the fROIs and the responses were estimated in the left-out run;

391   this procedure was repeated leaving out each run in turn; the response estimates were

392   then averaged across runs to derive a single estimate per condition per fROI. This

393   procedure allows all of the data to be used while maintaining the independence between

394   the data used to define the fROIs and the data used to examine their responses (e.g.,

395   Nieto-Castañon & Fedorenko 2012).

396   For the articulation regions, we used functional parcels derived from a group-level

397   probabilistic activation overlap map for the contrast between the production of difficult-

398   to-articulate nonwords and fixation in 20 participants, as reported in Basilakos et al.

399   (2017). We focused on the regions in the premotor cortices bilaterally: a region in the left

400   precentral gyrus, and two regions in the right precentral gyrus (see Figure 3 in Basilakos

401   et al. 2017). To define individual articulation-responsive fROIs, these parcels were

402   masked with activation maps for a contrast between imitating nonwords (repeating the

403   nonword produced by the actress) and imitating hand actions (in Experiment 3b). The

404   responses were then extracted to nonwords, and the four action imitation conditions. As

405   with the analyses of the speech-responsive regions, to estimate the responses to the

406   nonwords and hand action imitation conditions, an across-runs cross-validation procedure

407   was used so that the data to define the fROIs and estimate their responses were

18

408   independent (e.g., Kriegeskorte et al. 2011).

409

410   **Analyses.** In the critical analyses that examined the responses of the language fROIs to

411   the different action observation / imitation conditions, we used two-tailed t-tests to

412   compare the responses to each action condition against i) the low-level fixation baseline,

413   ii) nonword processing, which serves as the control condition in the language localizer,

414   and iii) sentence comprehension. The resulting p values were corrected for the number of

415   language fROIs within each experiment (i.e., 6), using the False Discovery Rate

416   correction (Benjamini and Yekutieli 2001). If language comprehension and action

417   observation / imitation share computational demands, then the action conditions should

418   elicit a response that is as strong as the sentence comprehension condition, or, at least,

419   reliably stronger than the nonword processing condition.

420

421   **Results**

422   *Behavioral data*

423   Overt behavioral responses were only collected in Experiment 1, where participants

424   watched videos and performed a one-back task on the action or the object in the video, as

425   described in Methods. Accuracies were high in both conditions, but slightly and reliably

426   higher for the actions condition than the objects condition (94.9% and 87.5%,

427   respectively; two-tailed $t(9)= 3.18$, $p < 0.05$). Further, as expected (given that actions

428   take time to unfold), participants were faster in the objects condition than the actions

429   condition (1.37s vs. 1.71s; two-tailed $t(9) = 6.05$, $p <= 0.0005$).

430   *Validation of the language fROIs*

19

431 Replicating previous work (Fedorenko et al. 2010; Fedorenko et al. 2011), the sentences

432 > nonwords effect was highly reliable in each of six fROIs both i) across the entire set of

433 participants (ts(76) > 10, ps < 0.0001), and ii) in each experiment individually

434 (Experiment 1: ts(9) > 4.43, ps < 0.001, Experiment 2: ts(53) > 8.39, ps < 0.0001,

435 Experiment 3a: ts(12) > 3.68, ps < 0.005, and Experiment 3b: ts(12) > 4.01, ps < 0.001).

436 Here, and in validating the MD fROIs, an across-runs cross-validation procedure,

437 described above, was used so that data used to define the fROIs were independent of the

438 data used to estimate the responses.

439 ***Reponses of the language fROIs to the action conditions***

440 The results are reported in Table 2 and Figure 2. Across experiments, none of the

441 language regions responded strongly and consistently to action observation or imitation.

442 In most fROIs, the action conditions failed to elicit a response above the fixation baseline

443 (except for Experiment 2, where both conditions elicited small but reliable above-

444 baseline responses in all language fROIs). Further, the response to the action

445 observation/imitation condition did not significantly differ from the nonword condition,

446 with the exception of the AngG fROI, which responded more strongly to some action

447 observation conditions than the nonword condition. Finally, again with the exception of

448 the AngG fROI, the response to the action observation/imitation condition was reliably

449 (in almost all cases, and always numerically) below that elicited by sentence

450 comprehension.

451 ***Experiment 1.*** When participants watched videos of a hand performing simple

452 manipulations of an object, there was no above-baseline response in any of the language

453 fROIs, regardless of whether participants were asked to focus on the objects (ts(9) < 1.5,

20

454 n.s.) or actions (ts(9) < 1.6, n.s.). Further, neither of the action conditions elicited a

455 response that was reliably greater than the nonword condition, whereas the sentence

456 condition elicited a reliably greater response than either of the two action conditions

457 (ts(9) > 2.75, ps < 0.05).

458 *Experiment 2.* In this experiment, every language fROI showed a reliably above-baseline

459 response to both the face action observation condition (ts(53) > 2.11, ps < 0.05) and the

460 body action observation condition (ts(53) > 2.86, ps < 0.01). However, in all fROIs

461 except for the AngG fROI, this response was i) not reliably higher than that elicited by

462 the nonword condition (ts(53) < 1.67, ps > 0.16), and ii) reliably lower than that elicited

463 by the sentence condition (ts(53) > 5.72, ps < 0.0001). In the AngG language fROI, both

464 action observation conditions elicited a response that was reliably stronger than that

465 elicited by the nonword condition and that did not differ from that elicited by the

466 sentence condition. We come back to the AngG fROI in the Discussion.

467 *Experiment 3a*. Similar to Experiment 1, there was no above-baseline response in the

468 language fROIs to any of the four conditions, with the exception of the AngG fROI and

469 the MFG fROI, which showed reliably above-baseline responses to hand action

470 observation (ts(12) > 2.82, ps < 0.05), but only the AngG fROI responded reliably more

471 strongly to hand action observation (and mouth action observation) than to nonwords

472 (ts(12) > 3.67, ps < 0.05); in all other fROIs none of the action observation conditions

473 produced a stronger response than nonwords. Finally, in all language fROIs, except for

474 the AngG fROI, the sentence condition elicited a reliably greater response than each of

475 the four action observation conditions (ts(12) > 3.30, ps < 0.01). In the AngG fROI, the

476 response to the action observation conditions did not reliably differ in magnitude from

21

477   the sentence condition.

478   ***Experiment 3b.*** In this experiment, where participants observed and imitated different

479   kinds of actions, there was no above-baseline responses except for the MFG fROI, which

480   responded reliably above baseline to the eye, mouth, and hand action conditions (ts(12) >

481   2.23, ps < 0.05), and marginally to the face action condition (t(12) = 3.09, p = 0.056).

482   However, these responses did not significantly differ from the response elicited by the

483   nonword condition (see Fedorenko et al. 2011, for a similar pattern of results with other

484   non-linguistic tasks). Further, the sentence condition elicited a reliably or marginally

485   greater response than each of the four action conditions in all language fROIs, except for

486   the AngG fROI and some frontal fROIs for some of the conditions (see Table 2 for

487   details).

488   ***Validation of the control, multiple demand (MD), fROIs***

489   Replicating previous work (Fedorenko et al. 2013; Blank et al. 2014), the hard > easy

490   spatial working memory effect was highly reliable in each of six fROIs across

491   participants with 2 runs (ts(47) > 7.8, ps < 0.0001). Participants with 1 run only (n=18)

492   could not be included in this validation analysis because across-runs cross-validation

493   could not be performed; for those participants, we ensured that MD activations looked as

494   expected based on visual examination of whole-brain activation maps.

495   ***Responses of the control, multiple demand (MD), fROIs to the action conditions***

496   Unlike in the language fROIs, all action imitation conditions elicited reliably above-

497   baseline responses in almost all MD fROIs. Similarly, at least some action observation

498   conditions elicited reliable responses. The body action observation condition from

499   Experiment 2, and the eye and hand action observation conditions from Experiment 3a

22

500 elicited the strongest responses. Strong responses to eye movement observation and

501 imitation could be related to prior claims about the role of this fronto-parietal system in

502 saccades (e.g., Pierrot-Deseilligny et al. 2004).

503 ***Responses of speech perception and articulation regions to the action conditions***

504 As discussed at the beginning of the Methods section, we have here focused on high-level

505 language processing regions (e.g., Fedorenko et al. 2010), which plausibly store our

506 linguistic knowledge that we use to both interpret and generate meaningful utterances

507 (e.g., Menenti et al. 2011). These regions are distinct from lower-level speech perception

508 regions (e.g., Overath et al. 2015; Norman-Haignere et al. 2015) and from speech

509 articulation regions (e.g., Bohland and Guenther 2006; Flinker et al. 2015; Basilakos et al.

510 2017). Might some of this perceptual or motor speech machinery overlap with action

511 observation or imitation? Based on the available evidence, a tentative answer appears to

512 be 'no'. In particular, the superior temporal regions that respond robustly to speech show

513 some response during speech articulation (e.g., Hickok et al. 2009; Basilakos et al. 2017),

514 but respond very little when participants produce even actions that involve speech

515 articulators, i.e., non-speech oral-motor movements (Basilakos et al. 2017). To shed

516 further light on this question, we performed an additional analysis on data from

517 Experiment 3a. We used a contrast between listening to nonwords and hand action

518 observation to define speech-responsive regions within the superior temporal cortex, and

519 then examined the responses of those regions to nonwords and hand action observation

520 (in data not used for fROI definition), as well as to face, eye, and mouth action

521 observation conditions. As Figure 4a clearly shows, the four action observation

522 conditions fail to elicit above-baseline responses, suggesting that these regions do not

23

523   support action observation.

524   What about regions that support speech articulation? Basilakos et al. (2017) report

525   quite robust responses to the production of non-speech oral-motor movements in

526   premotor articulation regions. We performed an additional analysis on data from

527   Experiment 3b to examine the responses of those articulation regions to action imitation

528   more broadly. We used a contrast between imitating nonwords (repeating the nonword

529   produced by the actress) and hand actions to define articulation-responsive regions within

530   ventral premotor cortex, and then examined the responses of those regions to nonwords

531   and hand action imitation (in data not used for fROI definition), as well as to face, eye,

532   and mouth action imitation. As Figure 4b shows, the mouth action imitation condition

533   elicits as strong a response as, or a stronger response than, articulation, replicating

534   Basilakos et al. (2017). The face condition (which also includes some mouth movements)

535   also elicits a strong response. However, the hand and eye action imitation conditions

536   elicit much lower responses. This relative selectivity for speech and oral-motor/face

537   actions is in line with the idea that these regions contain a map of our articulatory

538   apparatus (e.g., Bouchard et al. 2013; Guenther 2016), arguing against broad engagement

539   in action imitation, as well as with prior findings of somatotopic organization in the

540   motor areas (e.g., Watkins et al. 2003; Pulvermuller et al. 2006; D'Ausilio et al. 2009;

541   Murakami et al. 2011).

542   Thus, similar to high-level language processing regions, speech perception

543   regions do not appear to support action observation, and speech articulation regions do

544   not appear to support action imitation.

545

24

Discussion

We asked whether any part(s) of the language network – a set of brain regions that support high-level language processing (e.g., Fedorenko et al. 2010; Fedorenko & Thompson-Schill, 2014) – respond to action observation and/or imitation. Neural machinery that supports both language processing and some aspects of action observation/imitation has been postulated based on two distinct ideas. First, inspired by the discovery of mirror neurons in macaques (Rizolatti et al. 1988), some have argued that manual actions served as a fundamental precursor to linguistic communication in the evolution of our species (e.g., Arbib 2005; but see e.g., Tomasello 2008; Corballis 2003, for arguments for gesture-based origins of language that do not hinge on the mirror-neuron-based theorizing). Second, some have postulated an amodal hierarchical processor in the left frontal cortex (in or near "Broca's area") that is hypothesized to support both language processing and action perception/planning (e.g., Tettamanti and Weniger 2006; Fiebach and Schuboltz 2006; Koechlin and Jubault 2006).

Across three experiments (77 participants, 90 scanning sessions), we examined neural responses of functionally defined language regions to a broad range of *action observation* conditions, including hand actions with (Experiment 1) or without (Experiment 3a) a manipulable object, but also actions that involve the face or face parts (Experiments 2 and 3a) and body parts other than the hands (Experiment 2). In the fourth experiment (13 participants), we further examined responses of language regions to *action imitation*, again involving different face and body parts.

The key result is that – with a single exception discussed below – none of the language regions responded strongly and consistently to action observation or imitation.

25

569   In most language regions, the action conditions did not elicit a response above the

570   fixation baseline, which suggests that the language regions are as active during action

571   observation/imitation as they are when we are looking at a blank screen. Further, in most

572   language regions, the response to the action observation/imitation conditions i) did not

573   significantly differ from the response elicited by the nonword condition (the control

574   condition in the language localizer task), and ii) was reliably lower than the response

575   elicited by the sentence condition. These results suggest that language regions are

576   selective for language processing, in line with earlier work that established selectivity for

577   language relative to arithmetic, executive processing, music perception, and social

578   cognition (e.g., Fedorenko and Varley 2016). This conclusion is also consistent with

579   lesion studies that have reported dissociations between linguistic deficits and deficits in

580   action observation/production (e.g., Sirigu et al. 1998), and with a recent fMRI study that

581   showed that the degree of lateralization for language appears to be unrelated to the degree

582   of lateralization for action observation (Häberling et al. 2016).

583   The only exception was the language fROI in the angular gyrus. This region

584   responded more strongly to some action observation conditions than to nonwords, and, in

585   some cases, the response to action observation was not significantly lower than the

586   response to sentences. Evidence is accumulating that this region differs functionally from

587   the rest of the language network. In particular, it shows relatively low functional

588   correlations with other language regions during naturalistic cognition (e.g., Blank et al.

589   2014), including when using dynamic network modeling (Chai et al. 2016), and it shows

590   lower correlations in effect sizes and lateralization (e.g., Mahowald and Fedorenko 2016).

591   It also differs from the other language regions in sensitivity to linguistic and non-

26

592   linguistic manipulations. For example, the AngG language fROI was the only region that

593   did not show sensitivity to syntactic complexity (Blank et al. 2016), and it was the only

594   region that did not respond more strongly to sentences than photographs matched for

595   semantic content (Amit et al. 2017). The latter result suggests that the AngG language

596   fROI may respond to visual stimuli in general, as opposed to action observation

597   specifically. However, the precise role of this region in human cognition remains to be

598   discovered. One current hypothesis (formulated not specifically about the language-

599   responsive portion of the angular gyrus, but about the broad anatomical area) is that it is

600   "involved in all aspects of semantic processing" and contributes to "behaviors requiring

601   fluent conceptual combination" (e.g., Binder et al. 2009; cf. Lambon Ralph et al. 2017).

602   We now touch on four theoretical issues that the current results bear on.

603

604   **Gestural origins of language**

605   Just because in the modern human brain, language processing and action observation

606   appear to recruit non-overlapping machinery does not imply that our linguistic

607   communication system could not have arisen from the manual modality. In fact, this

608   possibility is still perhaps the most plausible (e.g., Tomasello 2008; Corballis 2003; cf.

609   Slocombe 2015; Shepherd & Freiwald, 2018). However, once humans began to develop

610   an extensive set of vocal communication signals, they plausibly had to allocate some

611   portions of the association cortices – massively expanded in the human brain (e.g.,

612   Buckner and Krienen 2013) – to store these form-meaning mappings (see also Häberling

613   et al., 2016). Given the differences between linguistic and non-linguistic communication

614   signals (including both discrete, categorical speech-accompanying gestures and

27

615   continuous, mimetic facial expressions and body language) – it is perhaps to be expected

616   that these different forms of communication would recruit distinct cognitive (e.g.,

617   Goldin-Meadow and Brentari 2016; McNeill 1992) and neural (e.g., Häberling et al.

618   2016) machinery given the distinct computational demands they place on the mind and

619   brain. It is worth noting that a few prior neuroimaging studies have argued that gesture

620   processing does recruit the same brain regions as language comprehension (e.g.,

621   Villarreal et al. 2008; Xu et al. 2009; Enrici et al. 2011; Andric et al. 2013; Redcay et al.

622   2016; see Willems & Hagoort 2007; Marstaller & Burianová 2014; Yang et al. 2015, for

623   reviews). However, those studies typically used symbolic gestures, pantomime, or

624   "emblems" (e.g., wave, hold out hand for a shake, etc.). Given that such gestures are

625   clearly associated with particular meanings, their processing may lead to the activation of

626   the corresponding linguistic representations. Thus, the overlap may be explained by the

627   engagement of linguistic resources during the processing of gestures rather than by the

628   shared computational demands like communicative intent or abstract conceptual

629   processing.

630

631   **Amodal hierarchical processor in "Broca's area"?**

632   Although our action observation/imitation conditions did not include a manipulation of

633   hierarchical complexity, we would argue that – to the extent that an amodal hierarchical

634   processor exists in the human brain – it does not reside within the high-level language

635   network. We have previously made this argument based on non-overlap between

636   language processing and music perception (e.g., Fedorenko et al. 2011; Fedorenko et al.

637   2012c; Norman-Haignere et al. 2015). Music is another domain that has been argued to

28

638   recruit such an amodal hierarchical processor (e.g., Maess et al. 2001; Koelsch et al.

639   2002). However, as Fedorenko & Varley (2016) have argued, the most compelling

640   evidence for overlap comes from structure-violation paradigms, and in those paradigms,

641   violations of structure appear to elicit similar responses to those elicited by low-level

642   oddball manipulations (e.g., Corbetta and Shulman 2002) and plausibly arise within the

643   domain-general multiple demand (MD) network due to increased processing effort

644   associated with unexpected events (Duncan 2010, 2013). Similarly, some manipulations

645   of hierarchical complexity in the action domain (e.g., Koechlin and Jubault 2006)

646   plausibly engage parts of the MD network because more complex action plans are

647   associated with greater working memory and cognitive control demands. Although parts

648   of the MD system have been argued to be particularly sensitive to hierarchical demands

649   (e.g., Badre and D'Esposito 2007, 2009; Badre 2008) or to the level of abstractness of the

650   to-be-processed information (e.g., Koechlin et al. 2003; Koechlin and Summerfield 2007),

651   these proposals have not gone unchallenged (e.g., Crittenden and Duncan 2012;

652   Pischedda et al. 2017). Thus, whether an amodal hierarchical processor exists anywhere

653   in the human brain remains an open question, but to the extent that it does, it exists

654   outside the boundaries of the high-level language network.

655

656   **Possibly similar computations across domains in spite of non-overlapping brain**

657   **regions**

658   The existence of distinct brain regions that support information processing in different

659   domains – like language vs. action observation vs. action production – does not imply

660   that the basic computations (that operate over those domain-specific representations) are

29

661   different. In fact, neural circuits across the cortex share many core properties (e.g.,

662   Douglas et al. 1989; Douglas and Martin 2004; Harris and Shepherd 2015), suggesting

663   that the basic computations may be the same or similar across different cortical areas. It is

664   also easy to come up with intuitive-level descriptions of potential parallels between

665   domains. For example, in the domain of language, we have a large store of form-meaning

666   mappings and knowledge about the relationships among them. We can use this

667   knowledge to interpret linguistic signals, and to generate new utterances, by combining

668   these basic building blocks into sequences. In the domain of actions, we may have a

669   similar "vocabulary" of actions for each of our effectors associated with particular

670   contexts of use, and information about how these actions can be combined (e.g., Hommel

671   et al. 2001; Rosenbaum et al. 2001; Schack 2004). And we can refer to this stored

672   knowledge to interpret others' actions as well as generate our own action sequences as

673   needed for goal-directed behavior (e.g., Flash and Bizzi 2016). As we make progress in

674   developing fleshed-out mechanistic-level hypotheses about what actually goes on when

675   we understand and produce language, or as we perceive and generate motor actions, it is

676   important to keep in mind both that i) the linguistic and action/motor representations

677   appear to be stored in non-overlapping brain areas, but that ii) the computations may be

678   fundamentally similar between these (and possibly other domains of perception, action,

679   and cognition).

680

681   **(Ir)relevance of the current results to the embodiment debate**

682   As noted in the Introduction, the question investigated here – i.e., whether high-level

683   language processing brain regions are engaged when we observe or produce motor

684   actions – is distinct from the much-debated question of the *nature of our conceptual*

685   *representations*. In particular, for many years now, some have advocated an "embodied"

686   view of meanings whereby concepts are "grounded" in sensory-motor modalities (e.g.,

687   Barsalou et al. 2003; Tranel et al. 2003; Simmons et al. 2007). Embodiment proposals

688   vary widely in the scope of their claims (see Leshinskaya and Caramazza 2016, for a

689   recent review of the key issues in this debate), from a complete denial of the existence of

690   abstract / amodal conceptual representations (e.g., Barsalou et al. 2003; Barsalou 2008;

691   Pulvermüller and Fadiga 2010; cf. Caramazza et al. 1990) to more moderate positions

692   where abstract representations interact in some way with the sensory/motor ones (e.g.,

693   Meteyard et al. 2012). The reason that the work reported here might, on the surface,

694   appear to be relevant to the embodiment debate is that action verbs have received a lot of

695   attention in that literature (e.g., see Bedny and Caramazza 2011, for a review). However,

696   the link is superficial: whether or not sensory and/or motor brain regions are active (to

697   some extent) when we understand the meanings of verbs like "kick" or "punch" (e.g.,

698   Hauk et al. 2004) is orthogonal to the question of whether the regions of the language

699   network – that we *know* are engaged when we process word meanings (e.g., Fedorenko et

700   al., 2012b) – play a role in the processing or execution of motor actions. We here show

701   that the answer to the latter question is no.

702

703   Before concluding, it is worth noting that, in general, the construct of "actions" is

704   complex and heterogeneous, and different researchers have different notions and scope in

705   mind when they talk about "actions". The conditions we included in our study have

706   spanned goal-directed/transitive actions (e.g., manipulating an object in Experiment 1)

707 and intransitive ones (e.g., tapping a finger in Experiment 3a), as well as actions

708 performed by different effectors (hand, feet, eyes, mouth, face). There are plausibly other

709 dimensions of actions that affect their representation and processing (e.g., Tarhan &

710 Konkle, 2017): e.g., whether the action is directed toward an animate entity vs. an object,

711 whether the action has communicative intent, whether the action maps onto a linguistic

712 label, etc. Our study leaves open the possibility that actions that have some specific

713 property/-ies may elicit responses in the language cortex. However, to evaluate this

714 possibility, we need clear testable hypotheses that would formally specify the relevant

715 features of actions that may lead to the recruitment of the same machinery as language

716 comprehension.

717

718 To conclude, action observation and action imitation do not recruit the left-lateralized

719 high-level language processing network, providing further evidence for the selectivity of

720 this network for language processing (e.g., Fedorenko and Varley 2016). However, this

721 separability is still compatible with hypotheses about the gestural origins of human

722 language (e.g., Tomasello 2008; Corballis 2003). Further, given the general similarity of

723 neural circuits across the neocortex (e.g., Harris and Shepherd 2015), research in the

724 domains of action perception or motor control may inform our understanding of the

725 computations that support language comprehension and production, domains where we

726 don't have the luxury of animal models to richly characterize neural response properties

727 and their interactions.

728

729

32

730 **Acknowledgements**

731

742

743 **Conflict of interest**

744
745 The authors declare no competing financial interests.
746

747 **References**

748

749 **Amit E, Hoeflin C, Hamzah N, Fedorenko E.** An asymmetrical relationship between

750   verbal and visual thinking: Converging evidence from behavior and fMRI.

751   *NeuroImage* 152: 619-627, 2017.

752 **Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HBM, Zilles K**. Broca's

753   Region Revisited: Cytoarchitecture and Intersubject Variability. *J Comp Neurol*

754   412: 319-341, 1999.

755 **Amunts K, Lenzen M, Friederici AD, Schleicher A, Morosan P, Palomero-Gallagher**

756   **N, Zilles K**. Broca's Region: Novel Organizational Principles and Multiple

757   Receptor Mapping. *PLoS Biol* 8: e1000489, 2010.

758 **Andric M, Solodkin A, Buccino G, Goldin-Meadow S, Rizzolatti G, Small SL.** Brain

759   Function Overlaps When People Observe Emblems, Speech, and Grasping.

760   *Neuropsychologia* 51: 1619-1629, 2013.

761 **Arbib MA**. From monkey-like action recognition to human language: An evolutionary

762   framework for neurolinguistics. *Behav Brain Sci* 28: 105-167, 2005.

763 **Arbib MA.** Mirror system activity for action and language is embedded in the integration

764   of dorsal and ventral pathways. *Brain Lang* 112: 12-24, 2010.

765 **Badre D, D'Esposito M**. Functional Magnetic Resonance Imaging Evidence for a

766   Hierarchical Organization of the Prefrontal Cortex. *J Cognitive Neurosci* 19:

767   2082-2099, 2007.

768 **Badre D**. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal

769   lobes. *Trends Cogn Sci* 12: 193-200, 2008.

770 **Badre D, D'Esposito M**. Is the rostro-caudal axis of the frontal lobe hierarchical?. *Nat*

34

771   *Neurosci* 10: 659-669, 2009.

772   **Barsalou LW, Simmons WK, Barbey AK, Wilson CD.** Grounding conceptual

773   knowledge in modality-specific systems. *Trends Cogn Sci* 7: 84-91, 2003.

774   **Barsalou LW.** Grounded cognition. *Annu Rev Psychol* 59: 617-645, 2008.

775   **Basilakos A, Smith KG, Fillmore P, Fridriksson J, Fedorenko E.** Functional

776   Characterization of the Human Speech Articulation Network. *Cereb Cortex* 1-15,

777   2017.

778   **Bautista A, Wilson SM**. Neural responses to grammatically and lexically degraded

779   speech. *Lang Cogn Neurosci* 31: 567-574, 2016.

780   **Bedny M, Caramazza A**. Perception, action, and word meanings in the human brain: the

781   case from action verbs. *Ann NY Acad Sci* 1224: 81-95, 2011.

782   **Benjamini Y, Yekutieli D**. The Control of the False Discovery Rate in Multiple Testing

783   under Dependency. *Ann Stat* 29: 1165-1188, 2001.

784   **Biagi L, Cioni G, Fogassi L, Guzzetta A, Sgandurra G, Tosetti M**. Action observation

785   network in childhood: a comparative fMRI study with adults. *Dev Sci* 19: 1075-

786   1086, 2016.

787   **Binder JR, Desai RH, Graves WW, Conant LL**. Where Is the Semantic System? A

788   Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies.

789   *Cereb Cortex* 19: 2767-2796, 2009.

790   **Binkofski F, Amunts K, Stephan KM, Posse S, Schormann T, Freund HJ, Zilles K,**

791   **Seitz RJ**. Broca's Region Subserves Imagery of Motion: A Combined

792   Cytoarchitectonic and fMRI Study. *Hum Brain Mapp* 11: 273-285, 2000.

793   **Blank I, Kanwisher N, Fedorenko E**. A functional dissociation between language and

794   multiple-demand systems revealed in patterns of BOLD signal fluctuations. *J*

795   *Neurophysiol* 112: 1105-1118, 2014.

796   **Blank I, Balewski Z, Mahowald K, Fedorenko E**. Syntactic processing is distributed

797   across the language system. *NeuroImage* 127: 307-323, 2016.

798   **Bohland JW, Guenther FH**. An fMRI investigation of syllable sequence production.

799   *NeuroImage* 32: 821-841, 2006.

800   **Bouchard KE, Mesgarani N, Johnson K, Chang EF.** Functional organization of human

801   sensorimotor cortex for speech articulation. *Nature* 495: 327-332, 2013.

802   (Corrigendum. Nature 498: 526, 2013)

803   **Buckner RL, Krienen FM**. The evolution of distributed association networks in the

804   human brain. *Trends Cogn Sci* 17: 648-665, 2013.

805   **Caramazza A, Hillis AE, Rapp BC, Romani C.** The Multiple Semantics Hypothesis:

806   Multiple Confusions?. *Cogn Neuropsychol* 7: 161-189, 1990.

807   **Caspers S, Zilles K, Laird AG, Eickhoff SB**. ALE meta-analysis of action observation

808   and imitation in the human brain. *NeuroImage* 50: 1148-1167, 2010.

809   **Chai LR, Mattar MG, Blank IA, Fedorenko E, Bassett DS**. Functional Network

810   Dynamics of the Language System. *Cereb Cortex* 1-12, 2016.

811   **Clerget E, Winderickx A, Fadiga L, Olivier E**. Role of Broca's area in encoding

812   sequential human actions: a virtual lesion study. *NeuroReport* 20: 1496-1499,

813   2009.

814   **Corballis MC**. From hand to mouth: The origins of language. Princeton, NJ: Princeton,

815   2003.

816   **Corballis MC.** Mirror neurons and the evolution of language. *Brain Lang* 112: 25-35,

817   2010.

818   **Corbetta M, Shulman GL**. Control of goal-directed and stimulus-driven attention in the

819   brain. *Nat Neurosci* 3: 201-215, 2002.

820   **Crittenden BM, Duncan J**. Task Difficulty Manipulation Reveals Multiple Demand

821   Activity but no Frontal Lobe Hierarchy. *Cereb Cortex* 24: 532-540, 2012.

822   **Culham JC, Valyear KF**. Human parietal cortex in action. *Curr Opin Neurobiol* 16:

823   205-212, 2006.

824   **D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, Fadiga L.** The

825   Motor Somatotopy of Speech Perception. *Curr Biol* 19: 381-385, 2009.

826   **Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL,**

827   **Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ**. An automated

828   labeling system for subdividing the human cerebral cortex on MRI scans into

829   gyral based regions of interest. *NeuroImage* 31: 968-980, 2006.

830   **Dinstein I, Hasson U, Rubin N, Heeger DJ**. Brain Areas Selective for Both Observed

831   and Executed Movements. *J Neurophysiol* 98: 1415-1427, 2007.

832   **Douglas RJ, Martin KAC, Whitteridge D**. A Canonical Microcircuit for Neocortex.

833   *Neural Comput* 1: 480-488, 1989.

834   **Douglas RJ, Martin KAC**. Neuronal circuits in the neocortex. *Ann Rev Neurosci* 27:

835   419-451, 2004.

836   **Duncan J, Owen AM**. Common regions of the human frontal lobe recruited by diverse

837   cognitive demands. *Trends Neurosci* 23: 475-483, 2000.

838   **Duncan J**. The multiple-demand (MD) system of the primate brain: mental programs for

839   intelligent behaviour. *Trends Cogn Sci* 14: 172-179, 2010.

37

840 **Duncan J**. The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron*

841   80: 35-50, 2013.

842 **Enrici I, Adenzato M, Cappa S, Bara BG, Tettamanti M**. Intention Processing in

843   Communication: A Common Brain Network for Language and Gestures. *J*

844   *Cognitive Neurosci* 23: 2415-2431, 2011.

845 **Fazio P, Cantagallo A, Craighero L, D'Ausilio AD, Roy AC, Pozzo T, Calzolari F,**

846   **Granieri E, Fadiga L**. Encoding of human action in Broca's area. *Brain* 132:

847   1980-1988, 2009.

848 **Fedorenko E, Hsieh PJ, Nieto-Castañón A, Whitfield-Gabrieli S, Kanwisher N**. New

849   Method for fMRI Investigations of Language: Defining ROIs Functionally in

850   Individual Subjects. *J Neurophysiol* 104: 1177-1194, 2010.

851 **Fedorenko E, Behr MK, Kanwisher N**. Functional specificity for high-level linguistic

852   processing in the human brain. *P Natl Acad Sci USA* 108: 16428-16433, 2011.

853 **Fedorenko E, Duncan J, Kanwisher N**. Language-Selective and Domain-General

854   Regions Lie Side by Side within Broca's Area. *Curr Bio* 22: 2059-2062, 2012a.

855 **Fedorenko E, Nieto-Castañon A, Kanwisher N**. Lexical and syntactic representations

856   in the brain: An fMRI investigation with multi-voxel pattern analysis.

857   *Neuropsychologia* 50: 499-513, 2012b.

858 **Fedorenko E, McDermott JH, Norman-Haignere S, Kanwisher N**. Sensitivity to

859   musical structure in the human brain. *J Neurophysiol* 108: 3289-3300, 2012c.

860 **Fedorenko E, Duncan J, Kanwisher N**. Broad domain generality in focal regions of

861   frontal and parietal cortex. *P Natl Acad Sci USA* 110: 16616-16621, 2013.

862 **Fedorenko E**. The role of domain-general cognitive control in language comprehension.

863   *Front Psychol* 5:1-17, 2014.

864   **Fedorenko E, Thompson-Schill SL.** Reworking the language network. *Trends Cogn Sci*

865   18: 120-126, 2014.

866   **Fedorenko E, Varley R**. Language and thought are not the same thing: evidence from

867   neuroimaging and neurological patients. *Ann NY Acad Sci* 1369: 132-153, 2016.

868   **Fiebach CJ, Schubotz RI**. Dynamic anticipatory processing of hierarchical sequential

869   events: a common role for Broca's area and ventral premotor cortex across

870   domains?. *Cortex* 42: 499-502, 2006.

871   **Fischl B, Rajendran N, Busa E, Augustinack J, Hinds O, Yeo BTT, Mohlberg H,**

872   **Amunts K, Zilles K**. Cortical Folding Patterns and Predicting Cytoarchitecture.

873   *Cereb Cortex* 18: 1973-1980, 2008.

874   **Flash T, Bizzi E**. Cortical circuits and modules in movement generation: experiments

875   and theories. *Curr Opin Neurobiol* 41: 174-178, 2016.

876   **Flinker A, Korzeniewska A, Shestyuk AY, Franaszczuk PJ, Dronkers NF, Knight**

877   **RT, Crone NE**. Redefining the role of Broca's area in speech. *Proc Natl Acad Sci*

878   *USA* 112: 2871-2875, 2015.

879   **Frost MA, Goebel R**. Measuring structural-functional correspondence: Spatial

880   variability of specialised brain regions after macro-anatomical alignment.

881   *NeuroImage* 59: 1369-1381, 2012.

882   **Gallivan JP, Culham JC**. Neural coding within human brain areas involved in actions.

883   *Curr Opin Neurobiol* 33: 141-149, 2015.

884   **Goldin-Meadow S, Brentari D**. Gesture, sign, and language: The coming of age of sign

885   language and gesture studies. *Behav Brain Sci* 1-82, 2016.

39

886   **Guenther FH**. *Neural Control of Speech.* Cambridge, MA: MIT Press, 2016.

887   **Häberling IS, Corballis PM, Corballis MC.** Language, gesture, and handedness:

888   Evidence for independent lateralized networks. *Cortex* 82: 72-85, 2016.

889   **Harris KD, Shepherd GMG.** The neocortical circuit: themes and variations. *Nat*

890   *Neurosci* 16: 170-181, 2015.

891   **Hauk O, Johnsrude I, Pulvermüller F.** Somatotopic Representation of Action Words in

892   Human Motor and Premotor Cortex. *Neuron* 41: 301-307, 2004.

893   **Hickok G, Okada K, Serences JT**. Area Spt in the Human Planum Temporale Supports

894   Sensory-Motor Integration for Speech Processing. *J Neurophysiol* 101: 2725-

895   2732, 2009.

896   **Hickok G**. Eight Problems for the Mirror Neuron Theory of Action Understanding in

897   Monkeys and Humans. *J Cogn Neurosci* 21: 1229-1243, 2009.

898   **Higuchi S, Chaminade T, Imamizu H, Kawato M**. Shared neural correlates for

899   language and tool use in Broca's area. *NeuroReport* 20: 1376-1381, 2009.

900   **Hommel B, Müsseler J, Aschersleben G, Prinz W**. The Theory of Event Coding

901   (TEC): a framework for perception and action planning. *Behav Brain Sci* 24:

902   849-878, 2001.

903   **Hugdahl K, Raichle ME, Mitra A, Specht K**. On the existence of a generalized non-

904   specific task-dependent network. *Front Hum Neurosci* 9: 430, 2015.

905   **Johnson-Frey SH, Newman-Norlund R, Grafton ST.** A Distributed Left Hemisphere

906   Network Active During Planning of Everyday Tool Use Skills. *Cereb Cortex* 15:

907   681-695, 2005.

908   **Juch H, Zimine I, Seghier ML, Lazeyras F, Fasel JHD**. Anatomical variability of the

909   lateral frontal lobe surface: Implication for intersubject variability in language

910   neuroimaging. *NeuroImage* 24: 504-514, 2005.

911   **Julian JB, Fedorenko E, Webster J, Kanwisher N**. An algorithmic method for

912   functionally defining regions of interest in the ventral visual pathway.

913   *NeuroImage* 60: 2357-2364, 2012.

914   **Kimura D**. Acquisition of a motor skill after left-hemisphere damage. *Brain* 100: 527-

915   542, 1977.

916   **Kimura D, Battison R, Lubert B**. Impairment of Nonlinguistic Hand Movements in a

917   Deaf Aphasic. *Brain Lang* 3: 566-571, 1976.

918   **Koechlin E, Ody C, Kouneiher F**. The Architecture of Cognitive Control in the Human

919   Prefrontal Cortex. *Science* 302: 1181-1185, 2003.

920   **Koechlin E, Jubault T**. Broca's Area and the Hierarchical Organization of Human

921   Behavior. *Neuron* 50: 963-974, 2006.

922   **Koechlin E, Summerfield C**. An information theoretical approach to prefrontal

923   executive function. *Trends Cogn Sci* 11: 229-235, 2007.

924   **Koelsch S, Gunter TC, Cramon DY, Zysset S, Lohmann G, Friederici AD**. Bach

925   Speaks: A Cortical "Language-Network" Serves the Processing of Music.

926   *NeuroImage* 17: 956-966, 2002.

927   **Kriegeskorte N**. Pattern-information analysis: From stimulus decoding to computational-

928   model testing. *NeuroImage* 56: 411-421, 2011.

929   **Lambon Ralph MA, Jefferies E, Patterson K, Rogers TT**. The neural and

930   computational bases of semantic cognition. *Nature* 18: 42-55, 2017.

931   **Leshinskaya A, Caramazza A.** For a cognitive neuroscience of concepts: moving

41

932   beyond the grounding issue. *Psychon Bull Rev* 23: 991-1001, 2016.

933   **Lingnau A, Gesierich B, Caramazza A**. Asymmetric fMRI adaptation reveals no

934   evidence for mirror neurons in humans. *P Natl Acad Sci USA* 106: 9925-9930,

935   2009.

936   **Maess B, Koelsch S, Gunter TC, Friederici AD**. Musical syntax is processed in Broca's

937   area: an MEG study. *Nat Neurosci* 4: 540-545, 2001.

938   **Mahowald K, Fedorenko E**. Reliable individual-level neural markers of high-level

939   language processing: A necessary precursor for relating neural variability to

940   behavioral and genetic variability. *NeuroImage* 139: 74-93, 2016.

941   **Marstaller L, Burianová H.** The multisensory perception of co-speech gestures – A

942   review and meta-analysis of neuroimaging studies. *J Neurolinguist* 30: 69-77,

943   2014.

944   **McNeill D**. *Hand and Mind: What Gestures Reveal about Thought.* Chicago, OH:

945   University of Chicago Press, 1992.

946   **Meister IG, Iacoboni M**. No Language-Specific Activation during Linguistic Processing

947   of Observed Actions. *PLoS ONE* 2: e891, 2007.

948   **Menenti L, Gierhan SME, Segaert K, Hagoort P**. Shared Language: Overlap and

949   Segregation of the Neuronal Infrastructure for Speaking and Listening Revealed

950   by Functional MRI. *Psychol Sci* 22: 1173-1182, 2011.

951   **Meteyard L, Rodriguez Cuadrado S, Bahrami B, Vigliocco G**. Coming of age: a

952   review of embodiment and the neuroscience of semantics. *Cortex* 48: 788-804,

953   2012.

954   **Molenberghs P, Cunnington R, Mattingley JB**. Brain regions with mirror properties: A

42

955   meta-analysis of 125 human fMRI studies. *Neurosci Biobehav R* 36: 341-349,

956   2012.

957   **Monti MM, Osherson DN**. Logic, language and the brain. *Brain Res* 1428: 33-42, 2012.

958   **Mukamel R, Ekstrom AD, Kaplan J, Iacoboni M, Fried I**. Single-Neuron Responses

959   in Humans during Execution and Observation of Actions. *Curr Biol* 20: 750-756,

960   2010.

961   **Murakami T, Restle J, Ziemann U.** Observation-execution matching and action

962   inhibition in human primary motor cortex during viewing of speech-related lip

963   movements or listening to speech. *Nueropsychologia* 49: 2045-2054, 2011.

964   **Nieto-Castañón A, Fedorenko E**. Subject-specific functional localizers increase

965   sensitivity and functional resolution of multi-subject analyses. *NeuroImage* 63:

966   1646-1669, 2012.

967   **Norman-Haignere S, Kanwisher NG, McDermott JH**. Distinct Cortical Pathways for

968   Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron*

969   88: 1281-1296, 2015.

970   **Oldfield RC**. The Assessment and Analysis of Handedness: The Edinburgh Inventory.

971   *Neuropsychologia* 9: 97-113, 1971.

972   **Overath T, McDermott JH, Zarate JM, Poeppel D**. The cortical analysis of speech-

973   specific temporal structure revealed by responses to sound quilts. *Nat Neurosci*

974   18: 903-911, 2015.

975   **Papagno C, Della Sala S, Basso A.** Ideomotor apraxia without aphasia and aphasia

976   without apraxia: the anatomical support for a double dissociation. *J Neurol*

977   *Neurosur Ps* 56: 286-289, 1993.

43

978   **Petrides M, Pandya DN.** Distinct Parietal and Temporal Pathways to the Homologues of

979   Broca's Area in the Monkey. *PLoS Biol* 7: e1000170, 2009.

980   **Pierrot-Deseilligny C, Milea D, Müri RM**. Eye movement control by the cerebral

981   cortex. *Curr Opin Neurol* 17: 17-25, 2004.

982   **Pischedda D, Görgen K, Haynes JD, Reverberi C.** Neural Representations of

983   Hierarchical Rule Sets: The Human Control System Represents Rules Irrespective

984   of the Hierarchical Level to Which They Belong. *J Neurosci* 37: 12281-12296,

985   2017.

986   **Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N**. Differential selectivity

987   for dynamic versus static information in face-selective cortical regions.

988   *NeuroImage* 56: 2356-2363, 2011.

989   **Poldrack RA**. Can cognitive processes be inferred from neuroimaging data?. *Trends*

990   *Cogn Sci* 10: 59-63, 2006.

991   **Poldrack RA**. Inferring Mental States from Neuroimaging Data: From Reverse Inference

992   to Large-Scale Decoding. *Neuron* 72: 692-697, 2011.

993   **Pulvermüller F, Huss M, Kherif F, Moscoso del Prado Martin F, Hauk O, Shtyrov**

994   **Y.** Motor cortex maps articulatory features of speech sounds. *Proc Nat Acad Sci*

995   *USA* 103: 7865-7870, 2006.

996   **Pulvermüller F, Fadiga L.** Active perception: sensorimotor circuits as a cortical basis

997   for language. *Nat Neuro* 11: 351-360, 2010.

998   **Redcay E, Ludlum RS, Velnoskey KR, Kanwal S.** Communicative Signals Promote

999   Object Recognition Memory and Modulate the Right Posterior STS. *J Cognitive*

1000   *Neurosci* 28: 8-19, 2016.

1001 **Rizzolatti G, Camarda R, Fogassi L, Gentilucci M, Luppino G, Matelli M**. Function

1002   organization of inferior area 6 in the macaque monkey. *Exp Brain Res* 71: 491-

1003   507, 1988.

1004 **Rizzolatti G, Arbib MA**. Language within our grasp. *Trends Neurosci* 21: 188-194,

1005   1998.

1006 **Rosenbaum DA, Meulenbroek RJ, Vaughan J, Jansen C**. Posture-Based Motion

1007   Planning: Applications to Grasping. *Psychol Rev* 108: 709-734, 2001.

1008 **Saxe R, Brett M, Kanwisher N**. Divide and conquer: A defense of functional localizers.

1009   *NeuroImage* 30: 1088-1096, 2006.

1010 **Saygin AP, Wilson SM, Dronkers NF, Bates E**. Action comprehension in aphasia:

1011   Linguistic and non-linguistic deficits and their lesion correlates.

1012   *Neuropsychologia* 42: 1788-1804, 2004.

1013 **Schack T**. The cognitive architecture of complex movement. *Int J Sport Exercise Psych*

1014   2: 403-438, 2004.

1015 **Scott TL, Gallée J, Fedorenko E**. A new fun and robust version of an fMRI localizer for

1016   the frontotemporal language system. *Cogn Neurosci* 167-176, 2016.

1017 **Shepherd SV, Freiwald WA**. Functional Networks for Social Communication in the

1018   Macaque Monkey. *Neuron* 99: 413-420, 2018.

1019 **Simmons WK, Ramjee V, Beauchamp MS, McRae K, Martin A, Barsalou LW.** A

1020   common neural substrate for perceiving and knowing about color.

1021   *Neuropsychologia* 45: 2802-2810, 2007.

1022 **Sirigu A, Cohen L, Zalla T, Pradat-Diehl P, Van Eeckhout P, Grafman J, Agid Y**.

1023   Distinct Frontal Regions for Processing Sentence Syntax and Story Grammar.

1024   *Cortex* 34: 771-778, 1998.

1025   **Slocombe KE**. Vocal Communication in Primates. In: *Emerging Trends in the Social and*

1026   *Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*,

1027   edited by Kosslyn SM. Hoboken, NJ: John Wiley & Sons, Inc, 2015, p. 1-12.

1028   **Tahmasebi AM, Davis MH, Wild CJ, Rodd JM, Hakyemez H, Abolmaesumi P,**

1029   **Johnsrude IS**. Is the Link between Anatomical Structure and Function Equally

1030   Strong at All Cognitive Levels of Processing?. *Cereb Cortex* 22: 1593-1603, 2011.

1031   **Tarhan L, Konkle T**. Modeling the Neural Structure Underlying Human Action

1032   Perception [Online]. Conference on Cognitive Computational Neuroscience.

1033   https://www2.securecms.com/CCNeuro/docs-0/591b2fa768ed3fa74b57b899.pdf

1034   [2017].

1035   **Tettamanti M, Weniger D**. Broca's area: A supramodal hierarchical processor?. *Cortex*

1036   42: 491-494, 2006.

1037   **Thesen S, Heid O, Mueller E, Schad L**. Prospective Acquisition Correction for Head

1038   Motion With Image-Based Tracking for Real-Time fMRI. *Magn Reson Med* 44:

1039   457-465, 2000.

1040   **Tomaiuolo F, MacDonald JD, Caramanos Z, Posner G, Chiavaras M, Evans AC,**

1041   **Petrides M**. Morphology, morphometry and probability mapping of the pars

1042   opercularis of the inferior frontal gyrus: an in vivo MRI analysis. *Eur J Neurosci*

1043   11: 3033-3046, 1999.

1044   **Tomasello M**. *Origins of Human Communication*. Cambridge, MA: MIT Press, 2008.

1045   **Tranel D, Kemmerer D, Adolphs R, Damasio H, Damasio AR.** Neural correlates of

1046   conceptual knowledge for actions. *Cogn Neuropsychol* 20: 409-432, 2003.

46

1047 **Tremblay P, Dick AS.** Broca and Wernicke are dead, or moving past the classic model

1048   of language neurobiology. *Brain Lang* 162: 60-71, 2016.

1049 **Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix**

1050   **N, Mazoyer B, Joliot M**. Automated Anatomical Labeling of Activations in SPM

1051   Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject

1052   Brain. *NeuroImage* 15: 273-289, 2002.

1053 **Villarreal M, Fridman EA, Amengual A, Falasco G, Gercovich ER, Ulloa ER,**

1054   **Leiguarda RC.** The neural substrate of gesture recognition. *Neuropsychologia*

1055   46: 2371-2382, 2008.

1056 **Watkins KE, Strafella AP, Paus T.** Seeing and hearing speech excites the motor system

1057   involved in speech production. *Neuropsychologia* 41: 989-994, 2003.

1058 **Willems RM, Hagoort P.** Neural evidence for the interplay between language, gesture,

1059   and action: A review. *Brain Lang* 101: 278-289, 2007.

1060 **Willems RM, Van der Haegen L, Fisher SE, Francks C**. On the other hand: including

1061   left-handers in cognitive neuroscience and neurogenetics. *Nat Neuro* 15: 193-201,

1062   2014.

1063 **Xu J, Gannon PJ, Emmorey K, Smith JF, Braun AR.** Symbolic gestures and spoken

1064   language are processed by a common neural system. *P Natl Acad Sci USA* 106:

1065   20664-20669, 2009.

1066 **Yang J, Andric M, Mathew MM.** The neural basis of hand gesture comprehension: A

1067   meta-analysis of functional magnetic resonance imaging studies. *Neurosci*

1068   *Biobehav R* 57: 88-104, 2015.

1069
1070

47

1071 **TABLES**

1072 **Table 1.** Timing parameters for the different versions of the language localizer task.

| | Version | | |
	A	B	C
Number of participants	5	5	80
Task: Passive Reading or Memory?	M	M	PR
Words / nonwords per trial	8	12	12
Trial duration (ms)	4,800	6,000	6,000
Fixation	300	300	100
Presentation of each word / nonword	350	350	450
Fixation	---	---	500
Memory probe	1,350	1,000	---
Fixation	350	500	---
Trials per block	5	3	3
Block duration (s)	24	18	18
Blocks per condition (per run)	8	8	8
Conditions	Sentences	Sentences	Sentences
	Nonwords	Nonwords	Nonwords
Fixation block duration (s)	16	18	14
Number of fixation blocks	5	5	5
Total run time (s)	464	378	358
Number of runs	2	2	2

1073

1074 **Table 2.** Results for each experiment and condition for the six language fROIs. Note that
1075 although in Figure 2, we plot the language localizer responses across the entire set of
1076 participants in the current study, all the comparisons between the action conditions and
1077 the conditions of the language localizer experiment were performed within each
1078 experiment separately using two-tailed paired-samples *t*-tests. In columns 2 and 3,
1079 significance values indicate whether the action observation/imitation condition elicited a
1080 response reliably above the baseline (column 2) or reliably above the nonword condition
1081 (column 3). In column 4, significance values indicate whether the action
1082 observation/imitation condition elicited a response reliably below the sentence condition.
1083

Experiment	Condition	ROI	Action condition vs. fixation	Action condition vs. nonwords	Action condition vs. sentences
Expt. 1	hand action observation (attention to action)	*LIFGorb*	$t(9) = -2.29, p = 0.143$	$t(9) = -1.33, p = 0.357$	$t(9) = 4.18, p < 0.005$
		LIFG	$t(9) < |1|$, *n.s.*	$t(9) < |1|$, *n.s.*	$t(9) = 3.31, p < 0.05$
		LMFG	$t(9) = 1.59, p = 0.294$	$t(9) = -1.66, p = 0.357$	$t(9) = 5.67, p < 0.005$
		LAntTemp	$t(9) = -3.32, p = 0.053$	$t(9) < |1|$, *n.s.*	$t(9) = 3.88, p < 0.01$
		LPostTemp	$t(9) < |1|$, *n.s.*	$t(9) = -1.26, p = 0.357$	$t(9) = 4.65, p < 0.005$
		LAngG	$t(9) = -1.08, p = 0.463$	$t(9) = 2.33, p = 0.271$	$t(9) = 2.76, p < 0.05$
	hand action observation (attention to object)	*LIFGorb*	$t(9) = -4.35, p < 0.01$	$t(9) = -2.10, p = 0.130$	$t(9) = 5.42, p < 0.001$
		LIFG	$t(9) < |1|$, *n.s.*	$t(9) = -1.41, p = 0.288$	$t(9) = 4.24, p < 0.005$
		LMFG	$t(9) = 1.54, p = 0.234$	$t(9) = -2.45, p = 0.110$	$t(9) = 5.82, p < 0.001$
		LAntTemp	$t(9) = -5.82, p < 0.005$	$t(9) = -1.16, p = 0.291$	$t(9) = 4.72, p < 0.005$
		LPostTemp	$t(9) = -1.40, p = 0.234$	$t(9) = -2.52, p = 0.110$	$t(9) = 6.10, p < 0.001$

48

		LAngG	*t*(9) = -2.87, *p* < 0.05	*t*(9) = 1.12, *p* = 0.291	*t*(9) = 4.75, *p* < 0.005
Expt. 2	Face action observation	*LIFGorb*	*t*(53) = 3.30, *p* < 0.005	*t*(53) < |1|, *n.s.*	*t*(53) = 6.19, *p* < 0.0001
		LIFG	*t*(53) = 3.28, *p* < 0.005	*t*(53) < |1|, *n.s.*	*t*(53) = 7.98, *p* < 0.0001
		LMFG	*t*(53) = 2.12, *p* < 0.05	*t*(53) = -3.48, *p* < 0.005	*t*(53) = 9.36, *p* < 0.0001
		LAntTemp	*t*(53) = 3.41, *p* < 0.005	*t*(53) = 1.66, *p* = 0.157	*t*(53) = 8.89, *p* < 0.0001
		LPostTemp	*t*(53) = 4.14, *p* < 0.0005	*t*(53) = -1.65, *p* = 0.157	*t*(53) = 9.69, *p* < 0.0001
		LAngG	*t*(53) = 4.98, *p* < 0.0001	*t*(53) = 4.52, *p* < 0.0005	*t*(53) < |1|, *n.s.*
	Body action observation	*LIFGorb*	*t*(53) = 4.16, *p* < 0.0005	*t*(53) < |1|, *n.s.*	*t*(53) = 5.73, *p* < 0.0001
		LIFG	*t*(53) = 3.56, *p* < 0.005	*t*(53) < |1|, *n.s.*	*t*(53) = 7.76, *p* < 0.0001
		LMFG	*t*(53) = 3.46, *p* < 0.005	*t*(53) = -3.57, *p* < 0.005	*t*(53) = 10.50, *p* < 0.0001
		LAntTemp	*t*(53) = 2.87, *p* < 0.01	*t*(53) < |1|, *n.s.*	*t*(53) = 9.64, *p* < 0.0001
		LPostTemp	*t*(53) = 3.23, *p* < 0.005	*t*(53) = -2.22, *p* = 0.061	*t*(53) = 10.31, *p* < 0.0001
		LAngG	*t*(53) = 6.66, *p* < 0.0001	*t*(53) = 6.00, *p* < 0.0001	*t*(53) = -1.83, *p* = 0.073
Expt. 3a	Face action observation	*LIFGorb*	*t*(12) = 1.76, *p* = 0.156	*t*(12) < |1|, *n.s.*	*t*(12) = 4.74, *p* < 0.001
		LIFG	*t*(12) = 2.53, *p* = 0.146	*t*(12) < |1|, *n.s.*	*t*(12) = 3.31, *p* < 0.01
		LMFG	*t*(12) < |1|, *n.s.*	*t*(12) = -2.11, *p* = 0.169	*t*(12) = 5.17, *p* < 0.001
		LAntTemp	*t*(12) < |1|, *n.s.*	*t*(12) < |1|, *n.s.*	*t*(12) = 5.24, *p* < 0.001
		LPostTemp	*t*(12) = 1.97, *p* = 0.146	*t*(12) = -1.10, *p* = 0.582	*t*(12) = 4.78, *p* < 0.001
		LAngG	*t*(12) = 2.19, *p* = 0.146	*t*(12) = 2.65, *p* = 0.126	*t*(12) = 1.13, *p* = 0.282
	Eye action observation	*LIFGorb*	*t*(12) < |1|, *n.s.*	*t*(12) < |1|, *n.s.*	*t*(12) = 4.03, *p* < 0.005
		LIFG	*t*(12) = 2.25, *p* = 0.263	*t*(12) < |1|, *n.s.*	*t*(12) = 3.64, *p* < 0.005
		LMFG	*t*(12) = 1.40, *p* = 0.562	*t*(12) = -1.20, *p* = 0.509	*t*(12) = 4.33, *p* < 0.005
		LAntTemp	*t*(12) < |1|, *n.s.*	*t*(12) < |1|, *n.s.*	*t*(12) = 6.95, *p* < 0.0001
		LPostTemp	*t*(12) < |1|, *n.s.*	*t*(12) = -1.37, *p* = 0.509	*t*(12) = 5.20, *p* < 0.001
		LAngG	*t*(12) < |1|, *n.s.*	*t*(12) = 1.66, *p* = 0.509	*t*(12) = 1.62, *p* = 0.131
	Mouth action observation	*LIFGorb*	*t*(12) < |1|, *n.s.*	*t*(12) = -2.11, *p* = 0.113	*t*(12) = 5.72, *p* < 0.0005
		LIFG	*t*(12) = 1.29, *p* = 0.440	*t*(12) < |1|, *n.s.*	*t*(12) = 4.10, *p* < 0.005
		LMFG	*t*(12) < |1|, *n.s.*	*t*(12) = -2.54, *p* = 0.078	*t*(12) = 5.11, *p* < 0.001
		LAntTemp	*t*(12) < |1|, *n.s.*	*t*(12) < |1|, *n.s.*	*t*(12) = 6.93, *p* < 0.0001
		LPostTemp	*t*(12) = 1.77, *p* = 0.328	*t*(12) = -1.20, *p* = 0.383	*t*(12) = 4.69, *p* < 0.001
		LAngG	*t*(12) = 1.73, *p* = 0.328	*t*(12) = 3.68, *p* < 0.05	*t*(12) = 1.04, *p* = 0.318
	Hand action observation	*LIFGorb*	*t*(12) < |1|, *n.s.*	*t*(12) = -1.14, *p* = 0.353	*t*(12) = 4.80, *p* < 0.001
		LIFG	*t*(12) = 2.04, *p* = 0.127	*t*(12) < |1|, *n.s.*	*t*(12) = 4.24, *p* < 0.005
		LMFG	*t*(12) = 2.83, *p* < 0.05	*t*(12) = -1.81, *p* = 0.190	*t*(12) = 5.13, *p* < 0.0005
		LAntTemp	*t*(12) = -1.55, *p* = 0.220	*t*(12) = -1.10, *p* = 0.353	*t*(12) = 5.84, *p* < 0.0005
		LPostTemp	*t*(12) = 1.15, *p* = 0.326	*t*(12) = -2.92, *p* < 0.05	*t*(12) = 6.48, *p* < 0.0005
		LAngG	*t*(12) = 3.31, *p* < 0.05	*t*(12) = 4.38, *p* < 0.01	*t*(12) < |1|, *n.s.*
Expt. 3b	Face action imitation	*LIFGorb*	*t*(12) = 2.06, *p* = 0.123	*t*(12) < |1|, *n.s.*	*t*(12) = 2.06, *p* = 0.075
		LIFG	*t*(12) = 1.08, *p* = 0.455	*t*(12) < |1|, *n.s.*	*t*(12) = 3.27, *p* < 0.05
		LMFG	*t*(12) = 3.09, *p* = 0.056	*t*(12) = -2.11, *p* = 0.169	*t*(12) = 2.52, *p* < 0.05
		LAntTemp	*t*(12) < |1|, *n.s.*	*t*(12) < |1|, *n.s.*	*t*(12) = 4.62, *p* < 0.005
		LPostTemp	*t*(12) = 2.37, *p* = 0.107	*t*(12) = -1.10, *p* = 0.582	*t*(12) = 4.10, *p* < 0.005
		LAngG	*t*(12) < |1|, *n.s.*	*t*(12) = 2.65, *p* = 0.126	*t*(12) = 1.77, *p* = 0.101
	Eye action imitation	*LIFGorb*	*t*(12) = 1.26, *p* = 0.386	*t*(12) < |1|, *n.s.*	*t*(12) = 1.78, *p* = 0.101
		LIFG	*t*(12) = 1.03, *p* = 0.386	*t*(12) < |1|, *n.s.*	*t*(12) = 3.17, *p* < 0.05
		LMFG	*t*(12) = 3.39, *p* < 0.05	*t*(12) = -1.47, *p* = 0.335	*t*(12) = 2.25, *p* = 0.053
		LAntTemp	*t*(12) = -1.04, *p* = 0.386	*t*(12) = -2.17, *p* = 0.304	*t*(12) = 6.06, *p* < 0.0005
		LPostTemp	*t*(12) = 1.44, *p* = 0.386	*t*(12) = -1.7, *p* = 0.335	*t*(12) = 5.16, *p* < 0.001
		LAngG	*t*(12) < |1|, *n.s.*	*t*(12) < |1|, *n.s.*	*t*(12) = 3.23, *p* < 0.05
	Mouth action imitation	*LIFGorb*	*t*(12) < |1|, *n.s.*	*t*(12) < |1|, *n.s.*	*t*(12) = 2.84, *p* < 0.05
		LIFG	*t*(12) = 1.27, *p* = 0.274	*t*(12) < |1|, *n.s.*	*t*(12) = 3.01, *p* < 0.05
		LMFG	*t*(12) = 3.24, *p* < 0.05	*t*(12) < |1|, *n.s.*	*t*(12) = 2.52, *p* < 0.05
		LAntTemp	*t*(12) = 2.04, *p* = 0.144	*t*(12) < |1|, *n.s.*	*t*(12) = 7.24, *p* < 0.0001

	LPostTemp	$t(12) = 1.85, p = 0.144$	$t(12) < |1|$, *n.s.*	$t(12) = 5.10, p < 0.001$
	LAngG	$t(12) = 1.81, p = 0.144$	$t(12) = 1.89, p = 0.495$	$t(12) = 1.22, p = 0.247$
Hand action imitation	*LIFGorb*	$t(12) = 1.32, p = 0.319$	$t(12) < |1|$, *n.s.*	$t(12) = 1.52, p = 0.163$
	LIFG	$t(12) = 1.78, p = 0.263$	$t(12) = 1.15, p = 0.549$	$t(12) = 1.77, p = 0.163$
	LMFG	$t(12) = 3.80, p < 0.05$	$t(12) = 1.34, p = 0.549$	$t(12) = 1.49, p = 0.163$
	LAntTemp	$t(12) < |1|$, *n.s.*	$t(12) < |1|$, *n.s.*	$t(12) = 4.76, p < 0.005$
	LPostTemp	$t(12) = 1.62, p = 0.263$	$t(12) < |1|$, *n.s.*	$t(12) = 4.16, p < 0.005$
	LAngG	$t(12) < |1|$, *n.s.*	$t(12) = 1.23, p = 0.549$	$t(12) = 1.49, p = 0.163$

1084
1085 **Table 3.** Results for each experiment (for Experiments 2 and 3a/b; no MD localizer was
1086 included in Experiment 1) and condition for the six MD fROIs. Significance values
1087 indicate whether the action observation/imitation condition elicited a response reliably
1088 above the baseline.
1089
1090

Experiment	Condition	ROI	Action condition vs. fixation
Expt. 2	Face action observation	*LIFGop*	$t(54) < |1|$, *n.s.*
		RIFGop	$t(54) = 1.15, p = 0.521$
		LPrecG	$t(54) < |1|$, *n.s.*
		RPrecG	$t(54) = 2.25, p = 0.169$
		LParSup	$t(54) = -1.66, p = 0.449$
		RParSup	$t(54) = -1.28, p = 0.521$
	Body action observation	*LIFGop*	$t(54) = 1.58, p = 0.197$
		RIFGop	$t(54) = 1.89, p = 0.521$
		LPrecG	$t(54) = 3.06, p < 0.05$
		RPrecG	$t(54) = 4.51, p < 0.0005$
		LParSup	$t(54) = 4.82, p < 0.0005$
		RParSup	$t(54) = 4.22, p < 0.001$
Expt. 3a	Face action observation	*LIFGop*	$t(12) = 1.27, p = 0.516$
		RIFGop	$t(12) = 1.46, p = 0.514$
		LPrecG	$t(12) = 1.40, p = 0.514$
		RPrecG	$t(12) = 1.97, p = 0.514$
		LParSup	$t(12) = 1.36, p = 0.514$
		RParSup	$t(12) < |1|$, *n.s.*
	Eye action observation	*LIFGop*	$t(12) = 2.52, p = 0.081$
		RIFGop	$t(12) = 3.68, p < 0.05$
		LPrecG	$t(12) = 3.40, p < 0.05$
		RPrecG	$t(12) = 4.66, p < 0.01$
		LParSup	$t(12) = 3.15, p < 0.05$
		RParSup	$t(12) = 2.03, p = 0.141$
	Mouth action observation	*LIFGop*	$t(12) = 2.32, p = 0.349$
		RIFGop	$t(12) = 1.40, p = 0.507$
		LPrecG	$t(12) = 1.89, p = 0.499$
		RPrecG	$t(12) = 2.47, p = 0.349$
		LParSup	$t(12) = 1.37, p = 0.507$
		RParSup	$t(12) < |1|$, *n.s.*
	Hand action observation	*LIFGop*	$t(12) = 2.00, p = 0.207$
		RIFGop	$t(12) = 2.38, p = 0.126$
		LPrecG	$t(12) = 2.67, p = 0.091$
		RPrecG	$t(12) = 3.91, p < 0.05$
		LParSup	$t(12) = 2.80, p = 0.091$
		RParSup	$t(12) = 1.83, p = 0.236$

50

Downloaded from www.physiology.org/journal/jn by ${individualUser.givenNames} ${individualUser.surname} (147.143.060.110) on September 13, 2018.
Copyright © 2018 American Physiological Society. All rights reserved.

Action response in language regions

	LPostTemp	$t(12) = 1.85, p = 0.144$	$t(12) < |1|$, *n.s.*	$t(12) = 5.10, p < 0.001$
	LAngG	$t(12) = 1.81, p = 0.144$	$t(12) = 1.89, p = 0.495$	$t(12) = 1.22, p = 0.247$
Hand action imitation	*LIFGorb*	$t(12) = 1.32, p = 0.319$	$t(12) < |1|$, *n.s.*	$t(12) = 1.52, p = 0.163$
	LIFG	$t(12) = 1.78, p = 0.263$	$t(12) = 1.15, p = 0.549$	$t(12) = 1.77, p = 0.163$
	LMFG	$t(12) = 3.80, p < 0.05$	$t(12) = 1.34, p = 0.549$	$t(12) = 1.49, p = 0.163$
	LAntTemp	$t(12) < |1|$, *n.s.*	$t(12) < |1|$, *n.s.*	$t(12) = 4.76, p < 0.005$
	LPostTemp	$t(12) = 1.62, p = 0.263$	$t(12) < |1|$, *n.s.*	$t(12) = 4.16, p < 0.005$
	LAngG	$t(12) < |1|$, *n.s.*	$t(12) = 1.23, p = 0.549$	$t(12) = 1.49, p = 0.163$

1084
1085 **Table 3.** Results for each experiment (for Experiments 2 and 3a/b; no MD localizer was
1086 included in Experiment 1) and condition for the six MD fROIs. Significance values
1087 indicate whether the action observation/imitation condition elicited a response reliably
1088 above the baseline.
1089
1090

Experiment	Condition	ROI	Action condition vs. fixation
Expt. 2	Face action observation	*LIFGop*	$t(54) < |1|$, *n.s.*
		RIFGop	$t(54) = 1.15, p = 0.521$
		LPrecG	$t(54) < |1|$, *n.s.*
		RPrecG	$t(54) = 2.25, p = 0.169$
		LParSup	$t(54) = -1.66, p = 0.449$
		RParSup	$t(54) = -1.28, p = 0.521$
	Body action observation	*LIFGop*	$t(54) = 1.58, p = 0.197$
		RIFGop	$t(54) = 1.89, p = 0.521$
		LPrecG	$t(54) = 3.06, p < 0.05$
		RPrecG	$t(54) = 4.51, p < 0.0005$
		LParSup	$t(54) = 4.82, p < 0.0005$
		RParSup	$t(54) = 4.22, p < 0.001$
Expt. 3a	Face action observation	*LIFGop*	$t(12) = 1.27, p = 0.516$
		RIFGop	$t(12) = 1.46, p = 0.514$
		LPrecG	$t(12) = 1.40, p = 0.514$
		RPrecG	$t(12) = 1.97, p = 0.514$
		LParSup	$t(12) = 1.36, p = 0.514$
		RParSup	$t(12) < |1|$, *n.s.*
	Eye action observation	*LIFGop*	$t(12) = 2.52, p = 0.081$
		RIFGop	$t(12) = 3.68, p < 0.05$
		LPrecG	$t(12) = 3.40, p < 0.05$
		RPrecG	$t(12) = 4.66, p < 0.01$
		LParSup	$t(12) = 3.15, p < 0.05$
		RParSup	$t(12) = 2.03, p = 0.141$
	Mouth action observation	*LIFGop*	$t(12) = 2.32, p = 0.349$
		RIFGop	$t(12) = 1.40, p = 0.507$
		LPrecG	$t(12) = 1.89, p = 0.499$
		RPrecG	$t(12) = 2.47, p = 0.349$
		LParSup	$t(12) = 1.37, p = 0.507$
		RParSup	$t(12) < |1|$, *n.s.*
	Hand action observation	*LIFGop*	$t(12) = 2.00, p = 0.207$
		RIFGop	$t(12) = 2.38, p = 0.126$
		LPrecG	$t(12) = 2.67, p = 0.091$
		RPrecG	$t(12) = 3.91, p < 0.05$
		LParSup	$t(12) = 2.80, p = 0.091$
		RParSup	$t(12) = 1.83, p = 0.236$

50

Action response in language regions

Expt. 3b	Face action imitation	*LIFGop*	$t(12) = 2.87, p < 0.05$
		RIFGop	$t(12) = 3.23, p < 0.05$
		LPrecG	$t(12) = 4.50, p < 0.005$
		RPrecG	$t(12) = 7.56, p < 0.0005$
		LParSup	$t(12) = 5.23, p < 0.001$
		RParSup	$t(12) = 3.29, p < 0.05$
	Eye action imitation	*LIFGop*	$t(12) = 2.09, p = 0.117$
		RIFGop	$t(12) = 2.71, p < 0.05$
		LPrecG	$t(12) = 2.78, p < 0.05$
		RPrecG	$t(12) = 3.64, p < 0.05$
		LParSup	$t(12) = 3.05, p < 0.05$
		RParSup	$t(12) = 1.80, p = 0.164$
	Mouth action imitation	*LIFGop*	$t(12) = 3.97, p < 0.01$
		RIFGop	$t(12) = 3.26, p < 0.05$
		LPrecG	$t(12) = 4.43, p < 0.005$
		RPrecG	$t(12) = 4.69, p < 0.005$
		LParSup	$t(12) = 4.17, p < 0.005$
		RParSup	$t(12) = 2.05, p = 0.088$
	Hand action imitation	*LIFGop*	$t(12) = 3.38, p < 0.01$
		RIFGop	$t(12) = 4.44, p < 0.005$
		LPrecG	$t(12) = 4.50, p < 0.005$
		RPrecG	$t(12) = 4.83, p < 0.005$
		LParSup	$t(12) = 4.24, p < 0.005$
		RParSup	$t(12) = 4.50, p < 0.005$

1091
1092
1093
1094
1095

51

1096   **Figure legends**

1097

1098   **Figure 1**: Sample stimuli for each experiment and condition. *Experiment 1*. a. Example

1099   objects, grouped vertically by family. b. Example family of dissimilar actions.

1100   *Experiment 2.* c. Example body action stimuli. d. Two sample face action stimuli.

1101   *Experiment 3a/b*. e. Example face actions. f. Example eye actions. g. Example mouth

1102   actions. h. Example hand actions.

1103

1104   **Figure 2**: Response to the language localizer conditions (estimated in data not used for

1105   fROI definition, as described in Methods) and action conditions across experiments. Next

1106   to each bar graph, we show the language parcels used to constrain the selection of

1107   individual language fROIs; the individual fROIs constitute 10% of each parcel (see

1108   Methods for details). Error bars indicate standard errors of the mean over participants.

1109

1110   **Figure 3**: Responses in multiple-demand regions to the action conditions in Experiments

1111   2 and 3a/b. Next to each bar graph, we show the MD parcels used to constrain the

1112   selection of individual MD fROIs; the individual fROIs constitute 10% of each parcel

1113   (see Methods for details). Error bars indicate standard errors of the mean over

1114   participants.

1115

1116   **Figure 4**: a. Responses in speech-responsive functional regions of interest (fROIs) in the

1117   auditory cortex (defined by nonword perception > hand action observation contrast in

1118   Experiment 3a, see Methods for details) to the nonword condition and the four action

1119   observation conditions in Experiment 3a. Responses are estimated using data not used for

52

1120   fROI definition (see Methods). Error bars indicate standard error of the mean over

1121   participants. Abbreviations, from left to right: L/R PP – left/right planum polare, L/R PT

1122   – left/right planum temporale, L/R ASTG – left/right anterior superior temporal gyrus,

1123   L/R PSTG – left/right posterior superior temporal gyrus.

1124   b. Responses in articulation-responsive fROIs in the premotor cortex (defined by the

1125   nonword imitation > hand action imitation contrast in Experiment 3b, see Methods for

1126   details) to the nonword imitation condition and the four action imitation conditions in

1127   Experiment 3b. Responses are estimated using data not used for fROI definition (see

1128   Methods). Error bars indicate standard error of the mean over participants. Abbreviations,

1129   from left to right: LPrCG – left precentral gyrus, RiPrCG – right inferior precentral gyrus,

1130   RsPrCG – right superior precentral gyrus.

1131

53

LIFGorb

LIFG

LMFG

LAntTemp

LPostTemp

LAngG

Legend:
- Sentences
- Nonwords
- hand action observation (attention to action)
- hand action observation (attention to object)
- face action observation
- body action observation
- face action observation/imitation
- eye action observation/imitation
- mouth action observation/imitation
- hand action observation/imitation

LIFGop

RIFGop

LPrecG

RPrecG

LParSup

RParSup

Legend:
- face action observation
- body action observation
- face action observation/imitation
- eye action observation/imitation
- mouth action observation/imitation
- hand action observation/imitation

Expt. 3a: responses in speech-responsive fROIs

Expt. 3b: responses in articulation-responsive fROIs