

RUNNING HEAD: Influence of modality on conceptual representations

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

Evidence for shared conceptual representations for sign and speech

Samuel Evans^{*1,2}, Cathy Price³, Jörn Diedrichsen⁴, Eva Gutierrez-Sigut¹ & Mairéad MacSweeney¹

¹Institute of Cognitive Neuroscience, UCL
²Dept of Psychology, University of Westminster
³Wellcome Trust Centre for Neuroimaging, UCL
⁴Brain and Mind Institute, University of Western Ontario

Corresponding author:

****Dr Samuel Evans***
Institute of Cognitive Neuroscience, UCL
Alexandra House
17 Queen Square
London WC1N 3AZ

Email: S.Evans1@westminster.ac.uk

RUNNING HEAD: Influence of modality on conceptual representations

32 **Abstract**

33 Do different languages evoke different conceptual representations? If so, greatest
34 divergence might be expected between languages that differ most in structure, such
35 as sign and speech. Unlike speech bilinguals, hearing sign-speech bilinguals use
36 languages conveyed in different modalities. We used functional magnetic resonance
37 imaging and representational similarity analysis (RSA) to quantify the similarity of
38 semantic representations elicited by the same concepts presented in spoken British
39 English and British Sign Language in hearing, early sign-speech bilinguals. We
40 found shared representations for semantic categories in left posterior middle and
41 inferior temporal cortex. Despite shared category representations, the same spoken
42 words and signs did not elicit similar neural patterns. Thus, contrary to previous
43 univariate activation-based analyses of speech and sign perception, we show that
44 semantic representations evoked by speech and sign are only partially shared. This
45 demonstrates the unique perspective that sign languages and RSA provide in
46 understanding how language influences conceptual representation.

47

48

49

50

51

52

53

54

RUNNING HEAD: Influence of modality on conceptual representations

55 **Introduction**

56 Conceptual knowledge is fundamental to human cognition. Recent evidence
57 suggests that conceptual representations are flexible and contextually defined^{1,2}.
58 Does the language that we use influence the nature of stored conceptual
59 representations? If this is the case, we might predict that languages that differ most
60 in structure, such as sign and speech, would show the greatest divergence between
61 conceptual representations. Sign languages are visuo-spatial natural languages that
62 are distinct from surrounding spoken languages. Hearing people with signing deaf
63 parents are bilingual in sign and speech. These individuals offer a unique insight into
64 the influence of both modality and bilingualism on semantic processing.

65 Semantic cognition engages a distributed left lateralised fronto-temporo-
66 parietal network^{3,4}. Strong evidence for modality independent neural representations
67 comes from studies using multivariate cross-classification of functional Magnetic
68 Resonance Imaging (fMRI) data that show that neural patterns elicited by an item in
69 one modality (e.g., pictures) can predict patterns for the same item presented in a
70 different modality (e.g., spoken words). These studies have identified common
71 patterns within hearing participants for pictures, identifiable sounds and spoken and
72 written words in the inferior temporal, parietal and prefrontal cortex⁵⁻⁷. Data from
73 patients with semantic dementia also suggest an important role for the inferior
74 anterior temporal lobe in semantic cognition, as a modality independent “hub”².
75 However, studies of the influence of modality on semantic processing in hearing
76 participants might reflect the eliciting of common oral language representations via
77 visual and auditory stimuli^{8,9}. Therefore, contrasting representations evoked by sign
78 and speech in hearing sign-speech bilinguals, offers a stronger test of the influence

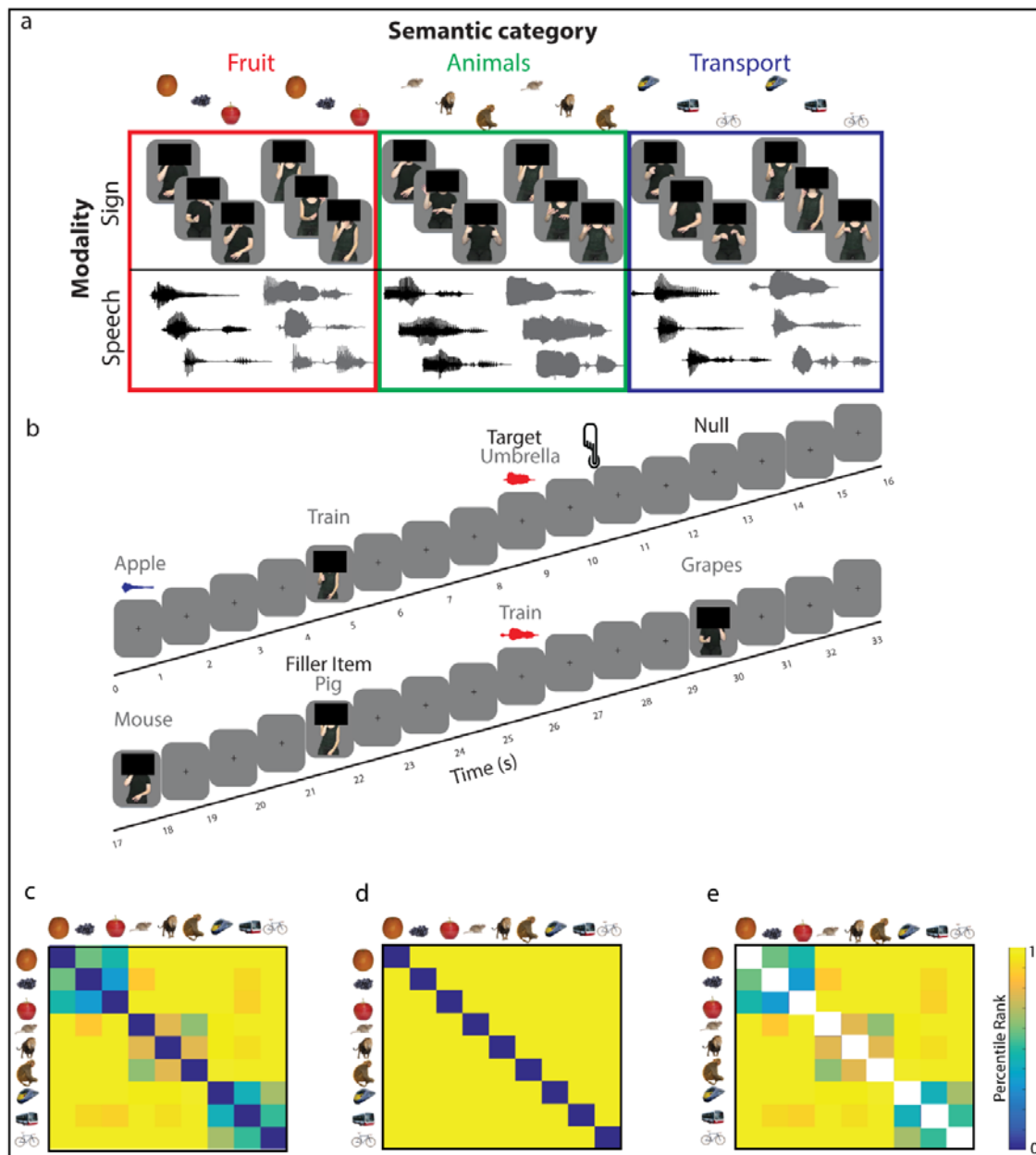
RUNNING HEAD: Influence of modality on conceptual representations

79 of modality on semantic processing, whilst also providing a unique perspective on
80 bilingualism.

81 How multiple languages are represented in a single brain is still not clear.
82 Evidence for shared representations comes from cross-linguistic priming¹⁰ and
83 stroop-type tasks¹¹ in spoken language bilinguals. However, evidence from word
84 association and translation tasks suggest different or only partially overlapping
85 semantic representations between languages^{12,13}. At the neural level, fMRI studies
86 show both common and language specific activity elicited by the different languages
87 of bilinguals^{14–18}. In these studies, the relative contribution of phonology, semantics
88 and syntactic processing has not been explicitly differentiated. Studies of bilinguals
89 to date have typically investigated across language representations ***within-***
90 ***modality***, e.g. from speech to speech, or text to text. Only one study has attempted
91 the stronger test of cross classifying between *both* language *and* modality. They
92 found it was not possible to cross-classify neural patterns for individual written and
93 heard words across different spoken languages¹⁹.

94 Sign and speech are conveyed in different modalities. Despite this, univariate
95 analyses of speech and sign perception reveal substantially overlapping brain
96 networks^{20–26}. However, to date, the similarity of neural patterns evoked by
97 individual signs and spoken words has not been quantified. Here, using
98 representational similarity analyses²⁷, we assess the evidence for shared and
99 language specific representations of individual conceptual items and semantic
100 categories, for speech and sign in hearing, early sign-speech bilinguals. Our
101 findings provide evidence for shared semantic representations at the level of
102 categories, but not for individual conceptual items. This suggests that visuo-spatial
103 languages and spoken languages evoke subtly different conceptual representations.

104



105

106 **Fig. 1. Stimuli, experimental design and semantic models. (Fig. 1a) Early sign-**
107 **speech bilinguals were presented with 9 conceptual items that belonged to 3**
108 **semantic categories: fruit, animals and transport. Items were presented as**
109 **signs and spoken words and were produced by male and female language**
110 **models. Video stills and oscillograms are shown for the signs and spoken**
111 **words respectively. Please note that the faces of the language models have**

RUNNING HEAD: Influence of modality on conceptual representations

112 **been obscured to comply with the policy of BioRxiv. Participants saw the**
113 **faces of the signers. (Fig. 1b) Within the scanner, participants attended to**
114 **speech and sign and pressed a button to identify items that were not in one of**
115 **the three target categories (e.g., umbrella). The dissimilarity between neural**
116 **patterns evoked by the signs and spoken words were tau-a correlated with**
117 **different theoretical models. These models included (Fig. 1c) a semantic**
118 **feature model derived from the CSLB concept property norms²⁸. The color bar**
119 **reflects the degree of semantic dissimilarity between items. This semantic**
120 **feature model can be decomposed into two independent components: (Fig. 1d)**
121 **An item-based dissimilarity model that predicts that each item is uniquely**
122 **represented, e.g., an ‘apple’ is more dissimilar to other items than to itself and**
123 **does not predict any broader semantic relatedness between items and (Fig 1e)**
124 **a category-based model in which the between-item similarities are predicted**
125 **by the semantic feature model, but where the within-item similarities are not**
126 **tested. White squares in this model indicate comparisons that were excluded.**

127

128 **RESULTS**

129 In the scanner, hearing early sign-speech bilinguals were presented with 9
130 conceptual items from the 3 semantic categories: fruit, animals or transport. Each
131 item was presented as a sign or as a spoken word and was produced by a male or a
132 female language model (Fig. 1a). Participants were instructed to press a button to
133 detect occasional items, 8% of the trials, that were not from one of the 3 target
134 categories (Fig. 1b). Performance in the scanner indicated that participants were
135 fully engaged with the semantic monitoring task (see Supplementary Information 1).
136 A univariate GLM analysis indicated that speech and sign language engaged similar
137 fronto-temporal networks, consistent with previous studies^{20–24} (see Supplementary
138 Information 2).

139

140 **Shared semantic representations for speech and sign**

141 Our criteria for identifying shared semantic representations for speech and
142 sign were as follows. First, using a searchlight analysis, we identified regions in
143 which there were reliably positive distances (see methods) between items **within-**
144 **modality** (e.g. averaging the speech-speech distances and the sign-sign distances).
145 We calculated distances only between items from the different language models
146 (e.g. different speakers and signers respectively) to exclude similarities driven by
147 low-level perceptual properties. In the identified regions, we then tested for **shared**
148 **semantic representations** applying the following criteria: (A) a significant fit to the
149 semantic feature model in the **within-modality** distances (e.g. both the across
150 speaker, speech-speech, and the across signer, sign-sign, distances) and (B) a
151 significant fit of the semantic feature model to the **across-modality** distances (e.g.
152 speech-sign and sign-speech distances). We also expected, (C) no evidence of a
153 difference in strength of fit to the semantic model between speech and sign, (D) no fit
154 to a model predicting greater distances between items from a different, as compared
155 to the same speaker, in the speech-speech distances, or from a different, as
156 compared to the same signer, in the sign-sign distances and (E) no fit to a model
157 predicting sensitivity to the iconicity of sign, a perceptual feature present in sign but
158 not speech.

159 Reliable within-modality distances were identified in six clusters (Fig. 2a): **(1)**
160 in bilateral V1-V3 and the LOC [-14 -96 10], **(2)** the right anterior superior temporal
161 gyrus [58 -4 -2], **(3)** the left anterior superior and middle temporal gyrus [-60 -10 -2],
162 **(4)** the right middle temporal gyrus and MT/V5 [52 -68 6], **(5)** the right insular [36 -12
163 14] and **(6)** the left posterior middle and inferior temporal gyrus (left pMTG/ITG) [-48 -
164 62 -6].

RUNNING HEAD: Influence of modality on conceptual representations

165 Only the response in the left posterior middle and inferior temporal gyri
166 (pMTG/ITG) cluster was consistent with shared semantic representations (see Fig.
167 2a cluster 6; Supplementary Information 3 for full details). In this cluster, there was a
168 significant fit to the (A) **within-modality** semantic feature model ($t(16) = 3.622$, $p =$
169 0.001 , $d_z = 0.879$, Fig 2d) and (B) **across-modality** semantic feature model ($t(16) =$
170 3.076 , $p = 0.004$, $d_z = 0.746$, Fig 2d). Whilst there was (C) no evidence for
171 differential sensitivity in the encoding of semantics for speech and sign ($t(16) =$
172 0.400 , $p = 0.694$, $d_z = 0.097$), (D) no sensitivity to the acoustic or visual features
173 associated with speaker (see model in Fig. 3e) or signer identity (see model in Fig.
174 4e), both $p_s > 0.063$, or (E) no influence of the iconicity structure of sign in the sign-
175 sign or across-modality distances, both $p_s > 0.106$ (see Supplementary Information 4
176 and Supplementary Fig. 2).

177 The fit of the semantic feature model (Fig. 1c) can be further decomposed into
178 item-based dissimilarity (Fig. 1d) and category-based dissimilarity (Fig. 1e). For
179 **within-modality** distances, the left pMTG/ITG region showed a significant fit to both
180 the semantic category ($t(16) = 1.980$, $p = 0.033$, $d_z = 0.480$) and item-based model (t
181 $(16) = 4.185$, $p = 3.50 \times 10^{-4}$, $d_z = 1.015$). The critical analyses **across-modality**,
182 indicated that the category-based model showed a significant fit to the data ($t(16) =$
183 2.509 , $p = 0.012$, $d_z = 0.608$), whereas the item-based model did not ($t(16) = 0.475$,
184 $p = 0.321$, $d_z = 0.115$). There was no evidence of a difference in the strength of fit to
185 the category model in the **within-modality** as compared to the **across-modality**
186 distances ($t(16) = 0.135$, $p = 0.894$, $d_z = 0.033$), suggesting that semantic
187 categories were represented equally robustly within- and across-modality. By
188 contrast, the item model was a significantly better fit to the **within-modality** than the
189 **across-modality** distances ($t(16) = 3.376$, $p = 0.004$, $d_z = 0.819$, Fig. 2f), providing

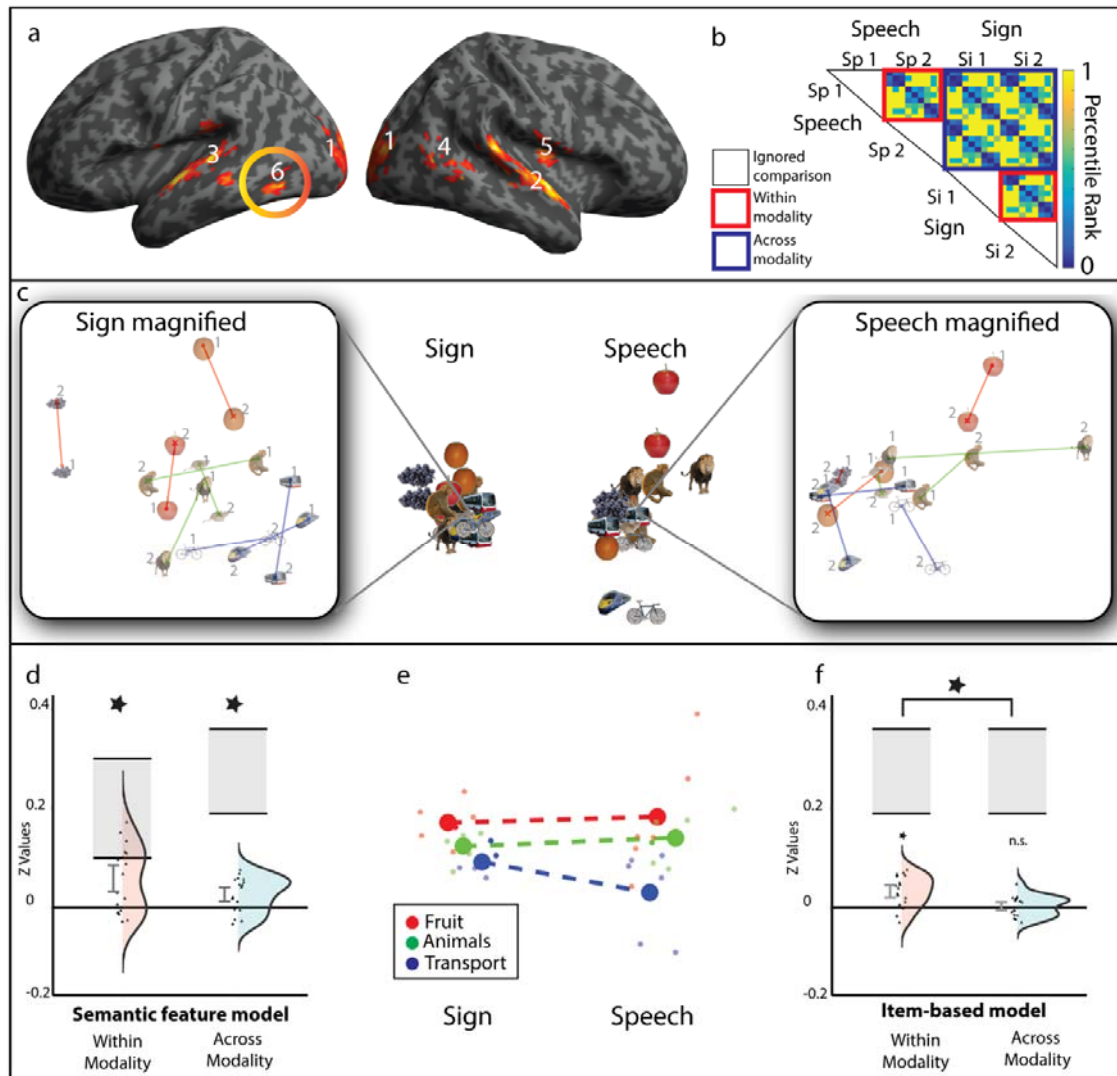
RUNNING HEAD: Influence of modality on conceptual representations

190 strong evidence that item-based representations are less robustly encoded across-
191 modality.

192 Together, these results suggest that semantic category structure drives the
193 commonality between activation patterns for sign and speech in left pMTG/ITG.
194 Indeed, this can be seen in the Multidimensional Scaling (MDS) solution (Fig. 2c)
195 used to visualise the similarity structure of the Representational Dissimilarity Matrix
196 (RDM). Fig. 2e illustrates the similar ordering of the category centroids both within
197 and across each modality.

198

RUNNING HEAD: Influence of modality on conceptual representations



199

200 **Fig. 2. Shared semantic representations for speech and sign. (Fig. 2a) A**
 201 **searchlight analysis identified brain regions containing positive *within-***
 202 ***modality* representational distances, thresholded at $p < 0.005$ peak level, FDR**
 203 **corrected at $q < 0.05$ at the cluster level. These regions are numbered**
 204 **according to the text in the results section. (Fig. 2b) Representational**
 205 **distances in these regions were Tau-a correlated with the semantic feature**
 206 **model within- and across-modality. The red boxes illustrate the within-**
 207 **modality distances, with the upper red box testing for abstracted speech**
 208 **representations (e.g. from speaker 1 to 2), and the lower red box testing for**
 209 **abstracted representations for sign (e.g. from signer 1 to 2). The blue box**
 210 **contains all across-language distances. Each 9x9 submatrix of dissimilarities**
 211 **is predicted from the semantic feature model (Fig. 1c). White boxes are**

RUNNING HEAD: Influence of modality on conceptual representations

212 **comparisons excluded from the analysis. The color bar reflects the predicted**
213 **strength of dissimilarity. Plots (Figs. 2c-f) show the response in cluster 6, the**
214 **left pMTG/ITG. (Fig. 2c) shows the non-metric MDS representation of the**
215 **response in left pMTG/ITG: the left panel shows within sign distances**
216 **magnified to make the representational structure clearer and the right panel**
217 **shows the equivalent speech representations. In these magnified images,**
218 **lines connect the same conceptual item produced by each speaker or signer,**
219 **marked as speaker/signer 1 or speaker/signer 2 on the figure. (Fig. 2d) In the**
220 **left pMTG/ITG, there was a significant fit to the semantic feature model in both**
221 **the within- and across-modality distances. Violin plots show distributions and**
222 **individual data points for the z transformed values, including the 90%**
223 **confidence interval and the noise ceiling (grey rectangle). The relative**
224 **contribution of item-based (Fig. 1d) and category-based (Fig. 1e) to this fit was**
225 **assessed. This showed there to be a significant fit to the category-based**
226 **model both within- and across-modality, without evidence of a difference in fit**
227 **when they were compared with one another. The MDS representation (Fig. 2e)**
228 **showing the mean centroid of each category within each modality for fruit**
229 **(red), animals (green), blue (transport), with dashed line connecting centroids**
230 **across-modality, highlights the within and across-modality category-based**
231 **dissimilarity. Plot (Fig. 2f) demonstrates that the item-based model was a**
232 **significant fit to the within-modality, but not across-modality distances, and**
233 **that the item-based model was a better fit to the within- as compared to**
234 **across-modality distances.**

235

236 **Modality specific representations**

237 Using a searchlight analysis, we tested for regions in which the average of the
238 speech-speech distances were greater than the sign-sign distances and vice versa.
239 This identified speech-specific and sign-specific processing regions. Within these
240 regions we tested for *modality specific semantic representations* evidenced by
241 (A) a fit to the semantic feature model (Fig. 1c) and (B) a fit to the semantic category

RUNNING HEAD: Influence of modality on conceptual representations

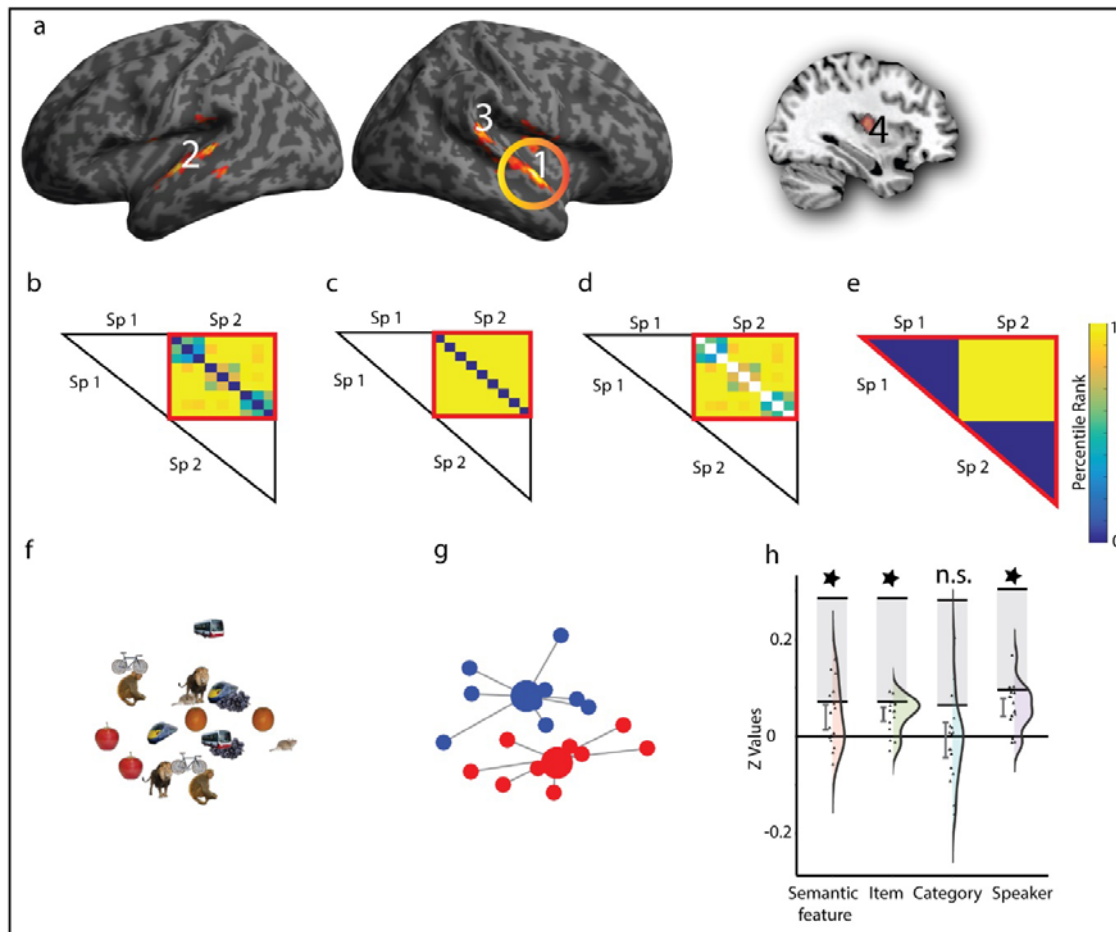
242 model (Fig. 1e) in the speech-speech or sign-sign distances for speech or sign
243 respectively and (C) no evidence of a fit to the speaker or signer identity model (see
244 the models in Fig. 3e and 4e).

245

246 *Speech specific responses*

247 For speech, the searchlight analysis revealed four clusters: **(1)** right anterior
248 STG extending to the temporal pole [58 -4 -2], **(2)** left anterior STG [-56 -8 2], **(3)**
249 right posterior STG/STS [58 -34 18] and **(4)** right putamen and insula [30 -10 10]
250 (see Fig. 3a). Within these regions, we tested for speech specific semantic
251 representations adjusting the critical alpha level to $p < 0.013$ to account for tests in
252 four clusters. In one of the four clusters, the right anterior STG [58 -4 -2] (Fig. 3a,
253 cluster 1), there was a significant fit to the semantic feature model ($t(16) = 2.529$, p
254 $= 0.011$, $d_z = 0.613$, see Fig 3b and Fig. 3h). This was driven by a fit to the item-level
255 model ($t(16) = 5.229$, $p = 4.14 \times 10^{-5}$, $d_z = 1.268$, see Fig. 3c and Fig. 3h). This
256 region was additionally sensitive to the acoustic differences between speakers ($t(16)$
257 $= 5.330$, $p = 3.39 \times 10^{-5}$, $d_z = 1.293$, see Fig. 3e and Fig. 3h) suggesting the
258 presence of speech form representations rather than speech selective semantic
259 representations (see Fig. 3f and Fig. 3g for MDS solution highlighting speaker-based
260 similarity). None of the four regions showed a response consistent with speech
261 specific semantic representations, as the category-based model (Fig. 3d) was not a
262 significant fit in any region (all p s > 0.110 , see fit to the speaker model in the right
263 STG in Fig. 3h).

264



265

266 **Fig. 3. Speech-specific neural responses. (Fig. 3a) A searchlight analysis**
267 **identified regions with greater representational distances for speech**
268 **compared to sign, thresholded at $p < 0.005$ peak level, FDR corrected at $q <$**
269 **0.05 at the cluster level. Clusters are numbered according to the text in the**
270 **results section. Models (Figs. 3b-e) show the within speech models that were**
271 **tested: (Fig. 3b) Within-speech semantic feature model, (Fig. 3c) Within-speech**
272 **item-based model, (Fig 3d) Within-speech category-based model and (Fig. 3e)**
273 **Between-speaker model. All models (Figs. 3b-d) test dissimilarities across**
274 **speaker (e.g. from speaker 1 to 2) in order to identify representations**
275 **abstracted from perceptual features. Color bar reflects predicted strength of**
276 **dissimilarity. White boxes are comparisons excluded from analysis. Plots**
277 **(Figs. 3f-h) show the response in cluster 1, the right anterior STG: (Fig. 3f)**
278 **Shows the non-metric MDS solution and (Fig. 3g) the same solution**
279 **highlighting speaker identity encoding. Large circles represent the centroids**

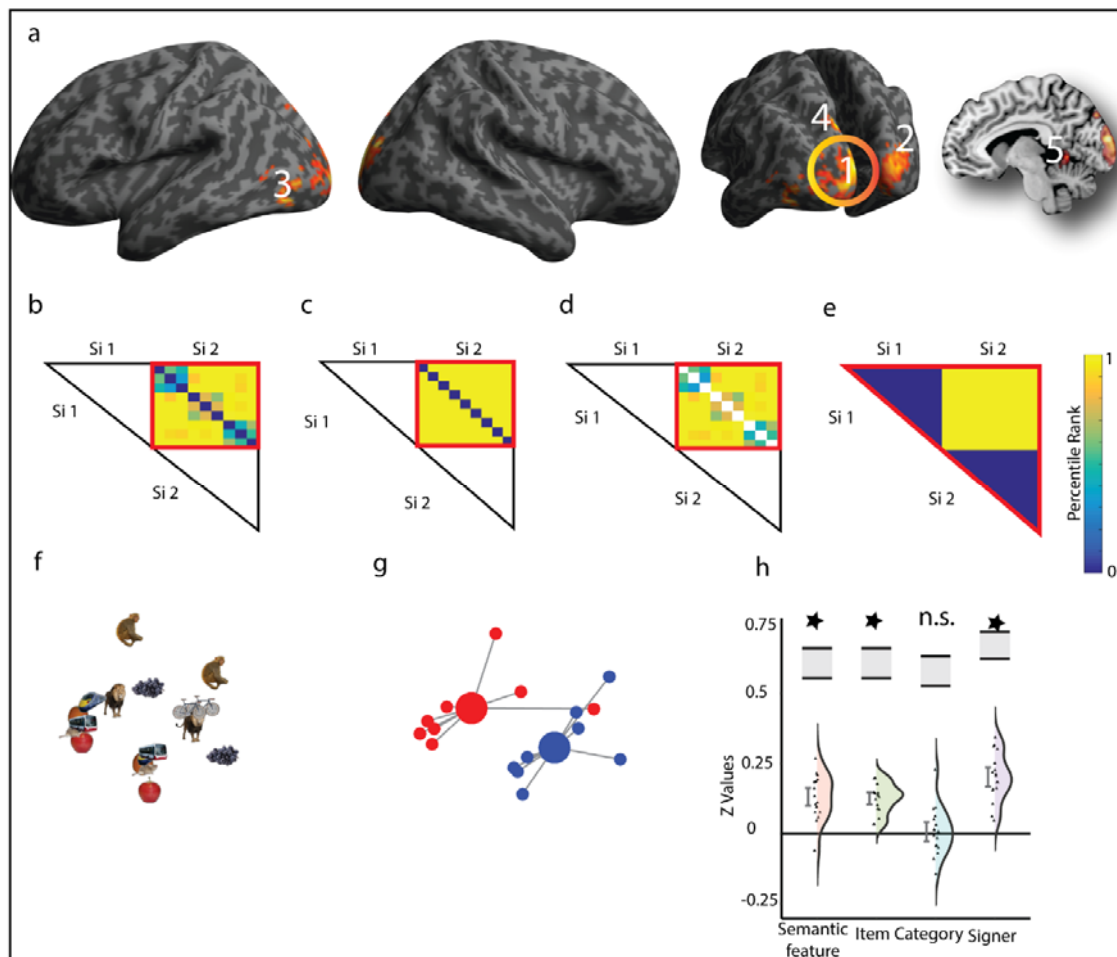
RUNNING HEAD: Influence of modality on conceptual representations

280 **for items from speaker 1 (red) and speaker 2 (blue). Smaller circles represent**
281 **the observed response for each item. Grey lines connect each item to**
282 **centroid. (Fig. 3h) Violin plots show model fits for z transformed values for**
283 **each model, with distributions and individual data points and 90% confidence**
284 **intervals and noise ceiling (grey box shown). This shows a significant fit to**
285 **the semantic feature model, driven by item-based rather than category-based**
286 **similarity structure and additional sensitivity to speaker identity, consistent**
287 **with abstract spoken word form representations rather than modality specific**
288 **semantic processing.**

289

290 *Sign specific responses*

291 Greater representational distances for sign than speech were identified in five
292 regions: **(1)** a cluster spreading across left V1-V3 [-6 -98 16], **(2)** a cluster within right
293 V1-V3 [22 -90 16], **(3)** a cluster in the left LOC and MT/V5 [-44 -80 -6], **(4)** left
294 superior occipital gyrus and superior parietal lobule [-10 -84 42] and **(5)** left lingual
295 gyrus spreading to the cerebellum [-4 -48 -8] (see Fig. 4a). Within these regions, we
296 tested for sign-specific semantic representations, adjusting the critical alpha level to
297 $p < 0.010$ to account for tests in five clusters. Analogous to the findings for speech,
298 the response in these regions was not consistent with sign-specific semantic
299 representations, as the category-based model was not a significant fit in any region
300 (all p s > 0.037). The response in clusters in the left V1-V3 and right V1-V3 cluster
301 were consistent with sign form representations characterised by a significant fit to the
302 semantic feature model (both p s $< 3.10 \times 10^{-5}$) but driven by item-based encoding
303 (p s $< 1.34 \times 10^{-7}$) with additional sensitivity to signer identity (both p s $< 3.29 \times 10^{-7}$,
304 see Fig. 4).



305

306 **Fig. 4. Sign specific neural responses. (Fig. 4a) A searchlight analysis**
307 **identified regions with greater representational distances for sign compared to**
308 **speech, thresholded at $p < 0.005$ peak level, FDR corrected at $q < 0.05$ at the**
309 **cluster level. Clusters are numbered according to the text in the results**
310 **section. Models (Figs. 4b-d) show the within signer models: (Fig. 4b) Within-signer**
311 **semantic feature model, (Fig. 4c) Within-signer item-based model, (Fig. 4d)**
312 **Within-signer category-based model and (Fig. 4e) Between-signer model.**
313 **Models (Figs. 4b-d) test dissimilarities across signer (e.g. from signer 1 to 2) to**
314 **identify representations abstracted from perceptual features. Color bar**
315 **reflects predicted strength of dissimilarity. White boxes are comparisons**
316 **excluded from analysis. Plots (Figs. 4f-h) show responses in cluster 1, the left**
317 **V1-V3. (Fig. 4f) Shows the non-metric MDS solution and (Fig. 4g) the same**
318 **solution highlighting the signer identity encoding in the left V1-V3 cluster.**
319 **Large circles represent the centroids for items from signer 1 (red) and signer 2**

RUNNING HEAD: Influence of modality on conceptual representations

320 **(blue). Smaller circles represent the observed response for each item. Grey**
321 **lines connect each item to centroid. (Fig. 4h) Violin plots show model fits for z**
322 **transformed values for each model fit, with distributions and individual data**
323 **points and 90% confidence intervals and noise ceiling (grey box shown). Plots**
324 **show a significant fit to the semantic feature model, driven by item-based**
325 **rather than category-based similarity structure and an additional sensitivity to**
326 **signer identity within the left V1-V3, consistent with abstract sign form**
327 **representations rather than modality specific semantic processing.**

328

329 **DISCUSSION** On the basis of univariate analyses of fMRI data it has been
330 assumed that the same underlying semantic representations support the perception
331 of spoken and signed languages²⁹. We tested this assumption, using RSA, to
332 quantify the similarity of neural patterns evoked by the same conceptual items
333 presented as BSL and spoken British English: two languages that differ in their
334 modality of expression. We tested for similarity at the level of individual items and
335 semantic categories. Shared category representations, that were abstracted from
336 surface acoustic and visual form, were found in the left pMTG/ITG. In this region,
337 both individual items and categories were encoded within-modality. Across-modality,
338 we found evidence for common coding of semantic categories. We did not detect
339 evidence of common item-level representations across modalities. Furthermore,
340 item-level encoding was significantly stronger within- as compared to across-
341 modality. In sign-specific and speech-specific areas, mainly in visual and auditory
342 primary and association cortices respectively, there was evidence for modality
343 specific item-based representations. In these regions, we did not see evidence for
344 category-based structure and the representations retained sensitivity to auditory and
345 visual features, suggestive of phonological word and sign form representations
346 rather than language specific semantic representations. Taken together, our data

RUNNING HEAD: Influence of modality on conceptual representations

347 are consistent with shared semantic representations between speech and sign, at
348 only a broad level of semantic specificity. In the following sections, we discuss the
349 implications of these findings.

350 **Shared semantic representations in pMTG/ITG** We identified shared
351 representations for semantic categories in sign and speech within the left pMTG/ITG.
352 This is consistent with studies showing common category representations for the
353 same items presented as pictures, environmental sounds, and spoken and written
354 words in this region^{5,7}. Indeed, activation of the left pMTG/ITG is associated with
355 the extraction of meaning from both the auditory and visual modalities. For example,
356 it is activated when reading words³⁰, in the perception of semantically ambiguous
357 speech³¹ and during sign language perception^{25,26,32}.

358 Common semantic coding for sign and speech was limited to category
359 representations and there was no evidence for direct correspondences between
360 individual spoken words and signs. Partially shared semantic representation
361 between languages is consistent with computational models of bilingualism, such as
362 the Distributed Feature Model³³. These models predict a single semantic store, in
363 which each language weights semantic features independently^{13,33,34}. The factors
364 contributing to differing weights between signed and spoken languages may be
365 greater than, and different to, those contributing to divergence between spoken
366 languages. Studies of spoken language processing show that lexical-semantic
367 access is affected by the phonological structure of the lexicon. For example, words
368 from dense phonological neighbourhoods activate semantic representations less
369 strongly³⁵ due to cascading activation between phonology and semantics³⁶. Indeed,
370 many computational models of speech processing do not make distinctions between
371 form and meaning³⁷. Similar architectures have been suggested for sign

RUNNING HEAD: Influence of modality on conceptual representations

372 processing³⁸. As natural languages, signed and spoken languages have very
373 different phonologies and phonological neighbourhoods. This might affect the
374 strength and structure of semantic activation within sign and speech lexicons, with
375 the possible result of reducing the commonality of conceptual representations
376 between the languages.

377 Another possibility is that the influence of greater iconicity found in sign
378 languages³⁹ may reduce the degree of similarity between semantic representations
379 of sign and speech. However, this is an unlikely explanation for the lack of item-
380 level correspondences between individual words and signs in the current dataset, as
381 we did not observe an effect of iconicity in the response in the left pMTG/ITG. There
382 are, however, more opaque form-meaning links that differ across speech and sign.
383 For example, the handshape “l” (extension of the little finger alone) denotes a
384 number of BSL signs that have negative connotations: bad, wrong, awful, poison⁴⁰.
385 Similarly, English words beginning with “gl” are often associated with light of low
386 intensity: gleam, glow, glint, glimmer, glint³⁹. Canonical signs can also carry
387 additional layers of meaning that allow communication of the size, location,
388 movement and other features of the referent; aspects of meaning that cannot be
389 communicated by the paralinguistic features of the voice. Again, these features may
390 fundamentally change the nature of semantic representation. These potential
391 explanations for the lack of item-level correspondences need to be tested in future.
392 For example, based on these findings, we might predict differences in the
393 representation of specific semantic categories, for example, representations for tools
394 might be expected to differ between unimodal (e.g. speech-speech) and bimodal
395 (e.g. sign-speech) bilinguals, on the basis that signs evoke greater specificity in the
396 semantic features associated with how they are handled.

RUNNING HEAD: Influence of modality on conceptual representations

397 An alternative explanation is that the absence of shared item-level
398 correspondences reflects the finer spatial scale of neural representations for
399 individual items which might be beyond the resolution of fMRI⁴¹. However, this would
400 seem unlikely given the identification of within-modality item-level encoding. Equally,
401 it might also reflect our methodological choices. We asked participants to monitor for
402 category rather than item-level distinctions⁴². We decided to use a category-based
403 task to maximise the likelihood of finding commonality between the languages, which
404 we assumed would be more robust at a broader level of semantic specificity.
405 Another possibility is that we did not have a high enough signal to noise ratio in
406 areas in which across-modality item level representations might be expected. A
407 posterior-anterior gradient of function has been suggested within the inferior
408 temporal cortex that reflects a wider-to-narrower window of semantic specificity^{2,43}.
409 The anterior inferior portion of the inferior temporal cortex is particularly susceptible
410 to signal drop out. Hence, the absence of shared item-level encoding might reflect
411 reduced signal quality in this region. However, tSNR maps for our data indicate
412 relatively good signal quality in most of the anterior inferior temporal cortex (see
413 Supplementary Information 3). Furthermore, drop out in the anterior inferior ATL
414 was similar to that found in the left pMTG/ITG and the superior ATL, regions in which
415 we found significant representational structure. We chose not to use a dual echo
416 sequence to mitigate against drop out⁴⁴, as our sequence was optimised for signal
417 quality in the posterior temporal cortex, the region most consistently activated by
418 both sign and speech in previous univariate studies. Future studies using dual echo
419 sequences and item-level discriminative tasks are necessary to exclude the
420 possibility that these methodological details obscured identification of item-level
421 correspondences in this study.

RUNNING HEAD: Influence of modality on conceptual representations

422 **Modality specific representations** Greater representational structure for
423 speech, than sign, was found in the bilateral superior temporal cortex and the right
424 insula. Within these regions, only a cluster in the right anterior superior temporal
425 cortex was a significant fit to the semantic model. This was shown to be driven by
426 the encoding of individual spoken words. A role for the anterior superior temporal
427 cortex in representing the identity of spoken words is consistent with studies in which
428 the intelligibility of speech has been parametrically varied or contrasted with non-
429 speech sounds^{45,46} and the suggestion that spoken word representations are
430 detected in the more superior portion of the ATL². This region was additionally
431 sensitive to speaker identity, suggesting that spoken word forms and speaker
432 characteristics are jointly encoded. This is consistent with a role for the right anterior
433 superior temporal cortex in representing speaker identity⁴⁷ and weak joint sensitivity
434 to spoken word and speaker identity in the right superior temporal cortex⁴⁸. The fact
435 that representations of spoken word forms were identified in the right, but not left
436 anterior STG, is unexpected. One possibility is that it is due to the greater
437 involvement of right hemisphere structures in language processing in early
438 bilinguals⁴⁹.

439 Regions containing greater representational structure for sign, than speech,
440 were found in the bilateral occipital cortices, as well as in the left superior parietal
441 lobule. This is consistent with the greater visual and body-space processing
442 demands of sign language perception²⁹ and the growing evidence for superior
443 parietal cortex involvement in sign perception and production⁵⁰. As for speech, a
444 subset of regions showing greater representational structure for sign than speech
445 showed a significant fit with the semantic model, and this was driven by item-level
446 encoding, consistent with visual sign form representations. Paralleling the findings

RUNNING HEAD: Influence of modality on conceptual representations

447 for speech, a number of these regions also exhibited a joint sensitivity to the identity
448 of the sign and the signer.

449 **Conclusions** For the first time, we quantified the similarity of neural
450 representations for the same conceptual items presented as sign and speech. We
451 found similarity between conceptual representations, at the category level, in the left
452 pMTG/ITG. We did not find evidence for regions in which there were direct one-to-
453 one mappings between individual spoken words and signs. This may suggest that
454 sign and speech share partially, but not fully, overlapping semantic representations.
455 This result is unexpected. Evidence to date has led researchers, including
456 ourselves, to propose extensive similarity in the neural processes underlying sign
457 and speech²⁹. Our findings suggest the need to rethink this assumption and
458 highlight the unique perspective that sign language can provide on language
459 processing and semantic representation more broadly.

460

461

462 **Online Methods**

463 **Participants** Ethical approval was granted by the UCL ethics committee. Data were
464 collected from 18 right handed early sign-speech bilinguals with no known
465 neurological, hearing or language learning impairments. One participant's data was
466 removed from the set due to an incidental finding, leaving a final data set of 17
467 participants (Mean age=33; range 20-52 years; female=12). Fifteen participants
468 learned British Sign Language (BSL) from a deaf parent and two from an older deaf
469 sibling. Two of the participants who learned sign language from a deaf parent did
470 not learn BSL from birth; one, learned AUSLAN from birth and learned BSL from the
471 age of twenty-one, the other, was exposed to another sign language from birth,

RUNNING HEAD: Influence of modality on conceptual representations

472 before learning BSL from 3 years of age. As a group the participants self-reported
473 excellent signing ability (mean = 6/7, SD= 0.86, range = 4-7).

474 **Stimuli** Stimuli consisted of nine core items for which neural responses were
475 analysed. Each core item was presented 48 times across the whole experiment, in
476 different modalities (sign/ speech) and by different models (male/ female) (see
477 'paradigm' for more details). These nine items belonged to three categories: fruit
478 (orange, grapes and apple), animals (mouse, lion and monkey) and transport (train,
479 bus and bicycle). Items within each category were similar and were distinct from
480 other categories on the basis of their semantic features, as evidenced by the CSLB
481 concept property norms²⁸ (see Fig. 1c). Items were chosen to ensure that the
482 categories were matched for age of acquisition (fruit M = 3.78; animals M = 4.52;
483 transport = 4.04), imageability (fruit M = 618; animals M = 610; transport M = 640),
484 familiarity (fruit M = 566; animals M = 521; transport M = 551) and the number of
485 syllables and phonemes in spoken English⁵¹⁻⁵⁴. In addition, we ensured that the
486 BSL equivalents of the spoken words were matched across category for handshape,
487 location, movement and handedness, and that iconicity⁵⁵ was similar across
488 categories (fruit M = 3.80; animals M = 3.92; transport M = 4.23; 1 low - 7 high
489 iconicity).

490 Speech samples were recorded by a male and female Southern British
491 English (SBE) speaker in an acoustically shielded booth with 16-bit quantisation and
492 a sampling rate of 22050 Hz using Adobe Audition. Spoken words were excised at
493 the zero crossing point. They were then filtered to account for the frequency
494 response of the Sensimetric headphones used in the scanner
495 (<http://www.sens.com/products/model-s14/>) and the overall amplitude was Root
496 Mean Square (RMS) equalised to ensure a similar perceived loudness (see Fig. 1a

RUNNING HEAD: Influence of modality on conceptual representations

497 for oscillograms). The mean duration of the auditory stimuli for the core items was
498 558ms (range = 323-865 ms), these sounds were similar in duration across semantic
499 categories (fruit M = 573 ms; animals M = 575 ms; transport M = 533 ms) and
500 gender of the speaker (male M = 557 ms; female M = 564 ms). The phonetic
501 distance between each of the spoken words was calculated using the Levenshtein
502 distance⁵⁶. This was achieved by calculating the number of phoneme insertions,
503 deletions and/or substitutions necessary to turn one word into the other, divided by
504 the number of phonemes in the longest word. The absolute value of the difference in
505 Levenshtein distance between each item was calculated. These distances did not
506 correlate with the semantic feature distances ($r = 0.063$, $n = 36$, $p = 0.713$), hence
507 semantic structure was not confounded with phonetic structure.

508 The BSL signs were all common variants in southern England as shown in the
509 BSL SignBank⁵⁷ (<http://bslsignbank.ucl.ac.uk/dictionary/>). Signs were recorded with
510 a Sony Handycam HDR-CX130 on a blue background by a male and a female deaf
511 native signer with a sampling rate of 50 fps and an aspect ratio of 1920x1080. The
512 blue background was keyed out and replaced with a dark grey background. Videos
513 were down-sampled to 30 frames per second and a resolution of 960 x 540 with
514 Adobe Premiere for presentation in the scanner. All signs were produced with
515 corresponding BSL mouthing. The signs were recorded in isolation such that the
516 hands returned to a neutral position resting on the knees between each sign. During
517 editing, the start and end-points of a sign were identified as a 'hold' (very brief pause
518 in movement of the hands) to remove the transitional movement into and out of the
519 neutral hands on the lap. Still frames of the hold points at the beginning and end of
520 each sign, with duration of 333ms, were inserted to ensure that the signs were easily
521 perceived in the scanner. The mean duration of the sign stimuli was 1107ms (range

RUNNING HEAD: Influence of modality on conceptual representations

522 = 867-1400ms). The signs were similar in duration as a function of semantic
523 category (fruit M = 1079ms; animals M = 1055ms; transport M = 1128ms) and
524 gender of the signer (male M = 1087ms; female M = 1086ms).

525 An iconicity dissimilarity measure⁵⁵ for the signs was calculated by taking the
526 absolute value of the difference between ratings of each item with every other.
527 These distances did not correlate with semantic feature similarity ($r = -0.126$, $n = 36$,
528 $p = 0.465$), hence semantic structure was not confounded with iconicity.

529 Participants were shown 36 additional items in the scanner to facilitate a
530 semantic monitoring task (see Fig. 1b) for which neural activity was not analysed.
531 The additional items consisted of 18 items from outside the categories of fruit, animal
532 and transport, e.g. buildings, clothes, furniture and tools, which were included as
533 target filler trials. Plus, an additional 18 non-target filler trials, 6 per category, of
534 other types of fruit, animals or transport that were included to reduce habituation to
535 the nine core items (see 'Paradigm' below for details of number of presentations).
536 Each individual filler item was produced by only one of the speakers or signers, with
537 the number of items from each speaker and signer balanced. The full set of stimuli
538 are available here: <https://osf.io/ek8ty/>.

539 Prior to scanning, participants were familiarised with the signs and spoken
540 words. Participants saw each sign stimulus and heard each word produced by both
541 sign and speech models and were required to name each item in spoken English.
542 They were shown all core items, target and non-target fillers. Sign recognition was
543 high (core items: mean = 17/18, min = 15/18, max = 18/18; filler items: mean =
544 32/36, min = 21/36, max = 35/36). On the very few occasions that participants
545 interpreted a sign as a non-intended English word, due to regional variations in

RUNNING HEAD: Influence of modality on conceptual representations

546 signs, participants were told the intended spoken label and asked to repeat it. They
547 were then retested on all the items in the experiment to ensure retention. Seventeen
548 out of 18 participants required one round of correction, the remaining participant
549 required a second round. Participants practiced a mock version of the within
550 scanner task on a laptop prior to scanning.

551 **Paradigm** In the scanner, participants were required to attend to the signed and
552 spoken stimuli and to press a button when they encountered an item from outside
553 the categories of fruit, animals or transport, e.g. a target filler item (see Fig. 1b). The
554 handedness of the button press was counterbalanced across participants.

555 Data were collected in 6 runs. In each run, each of the 9 core items were
556 presented twice in each of the following formats: sign and speech; male and female
557 model. Therefore each core item was presented 8 times in each run (2x2x2), with 72
558 core trials in total (9 items x 8 instances). Within each run, core items were
559 presented as two concatenated mini blocks of 36 trials. Within each mini block items
560 were randomised with the constraint that the same concept (e.g., 'orange') could not
561 be presented consecutively, regardless of modality, to reduce habituation.

562 In addition, in each run there were 6 target filler trials (non fruits, transport or
563 animals) for which participants were required to press a button and 6 non-target
564 fillers ('other' fruits, transport or animal items). The total number of trials was
565 balanced within run for modality (e.g. whether sign or speech) and language model
566 (e.g. speaker and signer). The filler trials (target and non-target fillers) were
567 interspersed within each run regularly but unpredictably. An additional, seven null
568 trials lasting 4 seconds were regularly but unpredictably interspersed within the each

RUNNING HEAD: Influence of modality on conceptual representations

569 run. During these trials a white fixation cross was presented on a grey background
570 in the absence of sound or additional visual stimulation for 4 seconds.

571 In summary, each of 6 runs consisted of 91 trials (72 core trials, 6 target filler
572 trials, 6 non-target filler trials, 7 null trials). The order of modality of presentation of
573 the items (speech/sign) was counter balanced across pairs of participants, such that
574 items presented as signs to participant 1 were presented as speech to participant 2,
575 and vice versa. Each stimulus was presented for its natural duration and was
576 followed by a fixation cross lasting 3 seconds, before the start of the next trial.

577 After scanning, participants provided iconicity ratings on the sign stimuli that
578 they had viewed in the scanner using the technique described by Vinson et al.⁵⁵.
579 They then took part in a multiple arrangement task in which they arranged pictures of
580 the core and filler items “based on their similarity” using a drag and drop interface⁵⁸.
581 The Euclidean distances derived from this arrangement correlated highly with the
582 CSLB concept property norms for the core items ($r = 0.904$, $n = 36$, $p = 4.42 \times 10^{-14}$),
583 suggesting that the semantic feature norms provided a good summary of the
584 semantic space of our participant group.

585 **Data Acquisition**

586 Data was acquired with a 3-Tesla scanner using a Magnetom TIM Trio
587 systems (Siemens Healthcare, Erlangen, Germany) with a 32 channel headcoil. A
588 2D epi sequence was used comprising forty 3mm thick slices using a continuous
589 ascending sequence (TR=2800ms, TA=2800ms, FA= 90°, TE=30ms, matrix size=
590 64x64, in-plane resolution: 3mm x 3mm, interslice gap = 1mm). Six runs of data
591 were acquired each lasting ~6-7 minutes with around 136 brain volumes collected
592 per run; the exact number of volumes was dependent on the stimuli included in each

RUNNING HEAD: Influence of modality on conceptual representations

593 run. EPI data collection lasted around 45 minutes. This was followed by a fieldmap,
594 acquired using a double-echo FLASH gradient echo sixty-four slice sequence
595 (TE1=10ms, TE2=12.46ms, in-plane view 192x192 mm, in-plane resolution: 3mm x
596 3mm, interslice gap = 1mm). At the end of the session a high-resolution T1 weighted
597 structural image was collected using a 3D Modified Driven Equilibrium Fourier
598 Transform (MDEFT) sequence (TR=1393ms, TE=2.48ms, FA= 16°, 176 slices, voxel
599 size = 1 × 1 × 1 mm).

600 In the scanner, stimuli were presented using the COGENT toolbox
601 (<http://www.vislab.ucl.ac.uk/cogent.php>) running in MATLAB. Auditory stimuli were
602 presented at the same comfortable listening level for all participants. Visual images
603 were presented using a JVC DLA-SX21 projector, with a screen resolution of
604 1024x768 and frame rate of 60Hz, using back projection onto a within bore screen at
605 a distance of 62cm from the participants' eyes.

606 **Univariate Analysis** Data were analysed using SPM12
607 (<http://www.fil.ion.ucl.ac.uk/spm/>). The first six images of each run were removed to
608 account for T1 equilibrium effects. The structural and functional images were
609 centred at the anterior commissure. Functional scans were slice time corrected to
610 the middle slice, realigned to the first image and unwarped using field maps. The
611 structural image was co-registered to the mean functional image. The parameters
612 derived from segmentation, using the revised SPM12 segmentation routines, were
613 applied to normalise the functional images that were re-sampled to 2x2x2mm. The
614 normalized images were then smoothed with a Gaussian kernel of 6-mm full-width
615 half maximum. Data were analyzed using a general linear model with a 360 second
616 high-pass filter and AR1 correction for auto-correlation. In the first level design
617 matrices, events were modelled with a canonical hemodynamic response function

RUNNING HEAD: Influence of modality on conceptual representations

618 marking the onset of the stimulus and duration in seconds. The design matrices
619 included a regressor for the onset of the speech trials, sign trials, filler target and
620 non-target trials in each modality (4 regressors), button presses when the target was
621 present in each modality (e.g. hits) (2 regressors) and button presses when the
622 target trials were absent for each modality (e.g. false alarms) (2 regressors), six
623 movement regressors of no interest and the session means. The rest condition
624 constituted an implicit baseline. Contrast images of [speech > rest] and [sign > rest]
625 were taken to the second level to conduct one sample t-tests.

626 **Representational similarity analysis (RSA)** At the first level, data were analysed
627 with SPM12. Analyses were conducted in native space. Images were slice time
628 corrected to the middle slice, realigned to the first image and unwrapped using
629 fieldmaps, but were not normalised or smoothed. The images were segmented,
630 using the revised SPM12 segmentation routine, to estimate the transformation from
631 native space to MNI space and vice versa. In the first level model in native space,
632 the two repetitions of each core item presented in each modality and by each
633 speaker and signer were modelled as a separate regressor (36 regressors: 9 core
634 items x 2 modalities x 2 language models). Additional regressors were included
635 modelling the onset of filler target and filler non-target trials for each modality (4
636 regressors), plus button presses when the target was present in each modality (e.g.
637 hits) (2 regressors) and button presses when the target trials were absent for each
638 modality (e.g. false alarms) (2 regressors). This constituted 42 regressors per run,
639 plus 6 motion parameter regressors and 6 session means. A high pass filter set at
640 360 seconds and AR(1) correction was applied. RSA analysis was conducted with
641 the latest version of the RSA toolbox (<https://github.com/rsagroup/rsatoolbox>)⁵⁹. The
642 representational distances estimated from the first level betas were used to calculate

RUNNING HEAD: Influence of modality on conceptual representations

643 the cross-validated Mahalanobis (crossnobis) distances using the RSA toolbox⁵⁹.
644 These crossnobis distances employ multivariate noise normalisation that down-
645 weight correlated noise across voxels, thereby increasing sensitivity to experimental
646 effects⁶⁰. The cross-validation across imaging runs ensures that the estimated
647 distances between neural patterns are not systematically biased by run-specific
648 noise, which allows us to test the distances directly against zero (as one would test
649 cross-validated classification accuracy against chance). Therefore, the crossnobis
650 distance provides a measurement on a ratio scale with an interpretable zero value
651 that reflects an absence of distance between items.

652 A volumetric searchlight analysis⁶¹ was conducted using a spherical 8mm
653 searchlight containing 65 voxels, consistent with the parameters used in previous
654 studies of language processing⁴⁸. In the searchlight analysis, the crossnobis
655 distance between each core stimulus and every other was calculated to generate a
656 Representational Dissimilarity Matrix (RDM) for every voxel and its surrounding
657 neighbourhood. The resulting RDM reflected sign-sign, speech-speech or speech-
658 sign distances, that constitute within and across-modality dissimilarities. In the
659 searchlight analyses, the average of speech-speech and sign-sign distances (e.g.
660 combined within-modality distances) and the average of the speech-speech and
661 sign-sign distances separately were returned to the voxel at the centre of each
662 sphere in three separate searchlight analyses. Within-modality distances were
663 calculated only between items from the different language models (e.g. different
664 speakers and signers respectively) to exclude similarities driven by low-level
665 perceptual properties. Each participants' native space whole brain searchlight map
666 was normalised to MNI space. These maps were inclusively masked with a >20%
667 probability grey matter mask, using the canonical MNI brain packaged with SPM12.

RUNNING HEAD: Influence of modality on conceptual representations

668 The resulting normalised, masked images were submitted to SPM12 for one sample
669 t-tests testing for greater than zero within-modality distances and paired t-tests
670 testing for differences between the speech-speech and sign-sign distances at the
671 second level. All statistical maps are presented at an uncorrected peak level
672 threshold of $p < 0.005$, FDR cluster corrected at $q < 0.05$ to identify regions of
673 interest for subsequent analysis.

674 The clusters identified from these analyses were used as Regions of Interest
675 (ROIs) in which to test theoretical models of brain function. Note that ROI analyses
676 are advised when testing special populations in which sample sizes are necessarily
677 restricted⁶². Using ROIs that contain reliable representational structure, e.g. greater
678 than zero distances, provides an additional protection against spurious distance-
679 model correlations in regions in which there is no reliable representational structure.
680 This approach is agnostic to the type of representational structure identified by the
681 searchlights ensuring that ROI selection and model validation are independent from
682 one another, and hence this does not represent “double dipping”⁶³.

683 As each cluster contains multiple RDMs, one for each searchlight contained
684 within the cluster, the RDMs were averaged, to provide a single representative RDM
685 for each cluster, and each participant. These distances were then used to test
686 hypothetical models of brain function (described below). The non-parametric Tau-a
687 correlation was used in preference to Pearson or Spearman correlation as the
688 models contained tied ranks⁵⁹. The resulting correlation coefficient was converted to
689 a Pearson’s r value, then to a Fisher-transformed Z value, to permit parametric
690 statistical analysis⁶⁴. Noise ceilings⁵⁹ were estimated within-modality and across-
691 modality separately as appropriate for each model. The lower bound was estimated
692 by calculating the mean z converted Tau-a correlation coefficient between each

RUNNING HEAD: Influence of modality on conceptual representations

693 participant's RDM and the average RDM for the group excluding that participant (e.g.
694 leaving one participant out). This is an estimate of the fit that should be achieved if
695 the theoretical model captures all systematic variation in the RDM across subjects in
696 this region. The upper bound was estimated by calculating the mean z converted,
697 Tau-a correlation between each participant's RDM and the average RDM for the
698 group including that participant. This value constitutes a theoretical maximum of the
699 best possible fit that can be achieved between the data and a model with this region.
700 These limits provide a benchmark against which to assess the quality of model fit as
701 they reflect the bounds of the best possible model fit that could be expected given
702 the noise in the data.

703 **Models**

704 A semantic model was tested using the CSLB concept property norms²⁸ (Fig.
705 1c). This kind of feature-based semantic model can account for the ability to
706 categorize by semantic group, e.g. a zebra is an animal, and to tell-apart unique
707 items, e.g. that a zebra differs from a horse. As such, the similarities expressed by
708 the model can be decomposed into two independent components. One, an **item-**
709 **based** model that predicts that each item is uniquely represented, e.g., an 'orange' is
710 more dissimilar to all other items than to itself, and does not predict any other
711 relatedness between items (Fig. 1d). The other, a model in which item-to-item
712 similarities are not tested, but category structure is predicted (Fig. 1e) – referred to
713 as a **category-based** model. An additional model testing for dissimilarities based on
714 speaker (Fig. 3e) and signer identity (Fig. 4e) was also tested, e.g. models predicting
715 trials from speaker/signer 1 to be more dissimilar than trials from speaker/signer 2,
716 and vice versa. The purpose of this model was to test for neural dissimilarities
717 based on lower level acoustic and visual features.

RUNNING HEAD: Influence of modality on conceptual representations

718 These models can be tested ***within-modality***, e.g. correlated within speech-
719 speech and sign-sign distances combined or separately, or ***across-modality***, e.g.
720 correlated with speech-sign distances. The testing of models using ***across-modality***
721 distances is equivalent to cross decoding representational structure between speech
722 and sign, positive evidence provides support for common representational structure
723 across languages⁶⁵. Note that we only test for ***across-modality*** semantic
724 representations in areas in which there is evidence of ***within-modality***
725 representational structure. As negative correlations are not plausible, greater than 0
726 model fits were assessed with one-tailed, one sample t-tests. Two-tailed paired t-
727 tests were used to assess differences in fit between models. Multidimensional
728 Scaling (MDS) was conducted to visualise the similarity structure of the RDMs by
729 calculating the averaged participant RDM and applying non-metric MDS, consistent
730 with the non-parametric correlational approach.

731

732

RUNNING HEAD: Influence of modality on conceptual representations

733 **Tables**

734 Table 1: MNI coordinates for RSA analyses – 3 local maxima more than 8 mm apart

Region	X	Y	Z	Extent	Z Value
Within-modality representational structure					
Right superior temporal gyrus	58	-4	-2	1545	5.283
Right inferior parietal lobule	64	-30	14		4.968
Right superior temporal gyrus	52	-2	-8		4.861
Left superior occipital gyrus	-14	-96	10	2629	4.677
Right superior occipital gyrus	14	-100	16		4.479
Right cuneus	6	-92	22		4.226
Left superior temporal gyrus	-60	-10	-2	1276	4.500
Left middle temporal gyrus	-64	-30	6		4.476
Left middle temporal gyrus	-64	-44	2		4.175
Left inferior temporal gyrus	-48	-62	-6	172	4.361
Left middle occipital gyrus	-42	-64	0		3.122
Right insula	36	-12	14	194	4.178
Right putamen	30	-8	10		4.160
Right middle temporal gyrus	52	-68	6	279	3.954
Right middle temporal gyrus	56	-48	0		3.748
Right middle temporal gyrus	54	-54	6		3.574
Greater representational structure for speech compared to sign					
Right superior temporal gyrus	58	-4	-2	754	4.877
Right superior temporal gyrus	52	0	-8		4.779

RUNNING HEAD: Influence of modality on conceptual representations

Right superior temporal gyrus	60	-12	4		3.590
Left superior temporal gyrus	-56	-8	2	743	4.484
Left superior temporal gyrus	-62	-30	10		4.253
Left superior temporal gyrus	-62	-2	0		3.720
Right Putamen	30	-10	10	146	4.364
Right Insular	40	-12	10		3.354
Right superior temporal gyrus	58	-34	18	285	4.160
Right superior temporal gyrus	66	-32	14		3.763
Right superior temporal gyrus	56	-26	0		3.722
Greater representational structure for sign compared to speech					
Left cuneus	-6	-98	16	1145	4.623
Left middle occipital gyrus	-12	-102	4		4.019
Left cuneus	-8	-94	28		3.830
Right superior occipital gyrus	22	-90	16	969	4.375
Right lingual gyrus	16	-84	-4		3.976
Right cuneus	16	-100	12		3.655
Left inferior occipital gyrus	-44	-80	-6	264	4.107
Left middle occipital gyrus	-50	-72	-2		3.937
Left middle occipital gyrus	-42	-80	4		3.449
Left cerebellum	-4	-48	-8	116	3.808
Left lingual gyrus	-10	-56	-2		3.767
Left cerebellum	-4	-50	0		3.102
Left superior occipital gyrus	-10	-84	42	127	3.781
Left superior occipital gyrus	-16	-78	40		3.396
Left superior parietal lobule	-26	-80	48		3.172

RUNNING HEAD: Influence of modality on conceptual representations

735 **DATA AVAILABILITY**

736 At the time of data collection participants did not consent to sharing their data via an
737 open repository. Therefore, the data of this study are not publicly available.

738 However, the data are available from the corresponding author upon request.

739

740 **ACKNOWLEDGEMENTS**

741 This research was funded by a Wellcome Trust Senior Research Fellowship
742 awarded to MM [100229/Z/12/Z]. CP is supported by a Wellcome Trust Principal
743 Research Fellowship [097720/Z/11/Z]. We would also like to acknowledge support
744 from an Economic and Social Research Council Research Centre Grant (Deafness
745 Cognition and Language Research Centre (DCAL) [RES-620-28-0002] and a
746 Wellcome Trust Centre Grant (203147/Z/16/Z). Thank you to Monika Grigorova and
747 Will Dawson for help in collecting this data.

748

749 **AUTHOR CONTRIBUTIONS**

750 S.E., M.M., J.D., C.P. & E.G. designed the study. S.E. collected the data. S.E., J.D.,
751 M.M. analysed the data. All authors contributed to writing the article.

752

753 **CONFLICTS OF INTEREST**

754 The authors declare no competing financial interests

755

756

757

758 **REFERENCES**

- 759 1. Kiefer, M. & Pulvermüller, F. Conceptual representations in mind and brain:
760 Theoretical developments, current evidence and future directions. *Cortex* **48**,
761 805–825 (2012).
- 762 2. Lambon Ralph, M., Jefferies, E., Patterson, K. & Rogers, T. T. The neural and
763 computational bases of semantic cognition. *Nat. Rev. Neurosci.* **18**, 42–55
764 (2016).
- 765 3. Binder, J. R., Desai, R. H., Graves, W. W. & Conant, L. L. Where Is the
766 Semantic System? A Critical Review and Meta-Analysis of 120 Functional
767 Neuroimaging Studies. *Cereb. Cortex* **19**, 2767–2796 (2009).
- 768 4. Price, C. J. A review and synthesis of the first 20years of PET and fMRI
769 studies of heard speech, spoken language and reading. *Neuroimage* **62**, 816–
770 847 (2012).
- 771 5. Fairhall, S. L. & Caramazza, A. Brain Regions That Represent Amodal
772 Conceptual Knowledge. *J. Neurosci.* **33**, 10552–10558 (2013).
- 773 6. Shinkareva, S. V, Malave, V. L., Mason, R. a, Mitchell, T. M. & Just, M. A.
774 Commonality of neural representations of words and pictures. *Neuroimage* **54**,
775 2418–25 (2011).
- 776 7. Simanova, I., Hagoort, P., Oostenveld, R. & van Gerven, M. A. J. Modality-
777 independent decoding of semantic information from the human brain. *Cereb.*
778 *Cortex* **24**, 426–34 (2014).
- 779 8. Araújo, S., Fernandes, T. & Huettig, F. Learning to read facilitates the retrieval
780 of phonological representations in rapid automatized naming: Evidence from

- 781 unschooled illiterate, ex-illiterate, and schooled literate adults. *Dev. Sci.*
782 e12783 (2018). doi:10.1111/desc.12783
- 783 9. Anthony, J. & Francis, D. Development of Phonological Awareness skill. *Curr.*
784 *Dir. Psychol. Sci.* **14**, 255–259 (2005).
- 785 10. Chen, H.-C. & Ng, M.-L. Semantic facilitation and translation priming effects in
786 Chinese-English bilinguals. *Mem. Cognit.* **17**, 454–462 (1989).
- 787 11. Preston, M. S. & Lambert, W. E. Interlingual interference in a bilingual version
788 of the stroop color-word task. *J. Verbal Learning Verbal Behav.* **8**, 295–301
789 (1969).
- 790 12. Kolers, P. A. Interlingual word associations. *J. Verbal Learning Verbal Behav.*
791 **2**, 291–300 (1963).
- 792 13. Duyck, W. & Brysbaert, M. Forward and backward number translation requires
793 conceptual mediation in both balanced and unbalanced bilinguals. *J. Exp.*
794 *Psychol. Hum. Percept. Perform.* **30**, 889–906 (2004).
- 795 14. Chee, M. W. L., Soon, C. S. & Lee, H. L. Common and segregated neuronal
796 networks for different languages revealed using functional magnetic resonance
797 adaptation. *J. Cogn. Neurosci.* **15**, 85–97 (2003).
- 798 15. Xu, M., Baldauf, D., Chang, C. Q., Desimone, R. & Tan, L. H. Distinct
799 Distributed patterns of neural activity are associated with two languages in the
800 bilingual brain. *Sci. Adv.* **3**, (2017).
- 801 16. Crinion, J. *et al.* Language control in the bilingual brain. *Science (80-)*. **312**,
802 1537–1540 (2006).
- 803 17. Buchweitz, A., Shinkareva, S. V., Mason, R. A., Mitchell, T. M. & Just, M. A.

- 804 Identifying bilingual semantic neural representations across languages. *Brain*
805 *Lang.* **120**, 282–289 (2012).
- 806 18. Correia, J. *et al.* Brain-Based Translation: fMRI Decoding of Spoken Words in
807 Bilinguals Reveals Language-Independent Semantic Representations in
808 Anterior Temporal Lobe. *J. Neurosci.* **34**, 332–338 (2014).
- 809 19. Van de Putte, E., De Baene, W., Price, C. J. & Duyck, W. “Neural overlap of L1
810 and L2 semantic representations across visual and auditory modalities: a
811 decoding approach”. *Neuropsychologia* **113**, 68–77 (2018).
- 812 20. MacSweeney, M. *et al.* Neural correlates of British sign language
813 comprehension: spatial processing demands of topographic language. *TL -*
814 *14. J. Cogn. Neurosci.* **14 VN-r**, 1064–1075 (2002).
- 815 21. Petitto, L. A. *et al.* Speech-like cerebral activity in profoundly deaf people
816 processing signed languages: Implications for the neural basis of human
817 language. *Proc. Natl. Acad. Sci.* **97**, 13961–13966 (2000).
- 818 22. Emmorey, K., McCullough, S. & Weisberg, J. Neural correlates of
819 fingerspelling, text, and sign processing in deaf American Sign Language–
820 English bilinguals. *Lang. Cogn. Neurosci.* **30**, 749–767 (2014).
- 821 23. Sakai, K. L., Tatsuno, Y., Suzuki, K., Kimura, H. & Ichida, Y. Sign and speech:
822 Amodal commonality in left hemisphere dominance for comprehension of
823 sentences. *Brain* **128**, 1407–1417 (2005).
- 824 24. Soderfeldt, B., Ronnberg, J. & Risberg, J. Regional Cerebral Blood Flow in
825 Sign Language Users. *Brain Lang.* **46**, 59–68 (1994).
- 826 25. MacSweeney, M. *et al.* Dissociating linguistic and nonlinguistic gestural

- 827 communication in the brain. *Neuroimage* **22**, 1605–18 (2004).
- 828 26. MacSweeney, M. *et al.* Lexical and sentential processing in British Sign
829 Language. *Hum. Brain Mapp.* **27**, 63–76 (2006).
- 830 27. Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis -
831 connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4
832 (2008).
- 833 28. Devereux, B. J., Tyler, L. K., Geertzen, J. & Randall, B. The Centre for
834 Speech, Language and the Brain (CSLB) concept property norms. *Behav.*
835 *Res. Methods* **46**, 1119–1127 (2014).
- 836 29. MacSweeney, M., Capek, C. M., Campbell, R. & Woll, B. The signing brain:
837 the neurobiology of sign language. *Trends Cogn. Sci.* **12**, 432–40 (2008).
- 838 30. Price, C. J., Wise, R. J. S. & Frackowiak, R. S. J. Demonstrating the implicit
839 processing of visually presented words and pseudowords. *Cereb. Cortex* **6**,
840 62–70 (1996).
- 841 31. Vitello, S., Warren, J. E., Devlin, J. T. & Rodd, J. M. Roles of frontal and
842 temporal regions in reinterpreting semantically ambiguous sentences. *Front.*
843 *Hum. Neurosci.* **8**, 1–14 (2014).
- 844 32. Macsweeney, M. *et al.* Neural systems underlying British Sign Language and
845 audio-visual English processing in native users. *Brain* **125**, 1583–1593 (2002).
- 846 33. de Groot, A. M. B. Determinants of Word Translation. *J. Exp. Psychol. Learn.*
847 *Mem. Cogn.* **18**, 1001–1018 (1992).
- 848 34. Van Hell, J. & De Groot, A. Conceptual representation in bilingual memory:
849 Effects of concreteness and cognate status in word association. *Biling. Lang.*

- 850 *Cogn. 1*, 193–211 (2003).
- 851 35. Marslen-Wilson, W. Activation, Competition, and Frequency in Lexical Access.
852 *Cogn. Model. Speech Process. Psycholinguist. Comput. Perspect.* 148–172
853 (1995).
- 854 36. Apfelbaum, K. S., Blumstein, S. E. & McMurray, B. Semantic priming is
855 affected by real-time phonological competition: Evidence for continuous
856 cascading systems. *Psychon. Bull. Rev.* **18**, 141–149 (2011).
- 857 37. Gaskell, M. G. & Marslen-Wilson, W. D. Integrating Form and Meaning: A
858 Distributed Model of Speech Perception. *Lang. Cogn. Process.* **12**, 613–656
859 (1997).
- 860 38. Caselli, N. K. & Cohen-Goldberg, A. M. Lexical access in sign language: A
861 computational model. *Front. Psychol.* **5**, 1–11 (2014).
- 862 39. Perniss, P., Thompson, R. L. & Vigliocco, G. Iconicity as a general property of
863 language: Evidence from spoken and signed languages. *Front. Psychol.* **1**, 1–
864 15 (2010).
- 865 40. Marshall, C., Rowley, K. & Atkinson, J. Modality-Dependent and -Independent
866 Factors in the Organisation of the Signed Language Lexicon: Insights From
867 Semantic and Phonological Fluency Tasks in BSL. *J. Psycholinguist. Res.* **43**,
868 587–610 (2013).
- 869 41. Gardumi, A. *et al.* The effect of spatial resolution on decoding accuracy in fMRI
870 multivariate pattern analysis. *Neuroimage* **132**, 32–42 (2016).
- 871 42. Bonte, M., Hausfeld, L., Scharke, W., Valente, G. & Formisano, E. Task-
872 dependent decoding of speaker and vowel identity from auditory cortical

- 873 response patterns. *J. Neurosci.* **34**, 4548–57 (2014).
- 874 43. Clarke, A. & Tyler, L. K. Understanding What We See: How We Derive
875 Meaning From Vision. *Trends Cogn. Sci.* **19**, 677–687 (2015).
- 876 44. Halai, A. D., Welbourne, S. R., Embleton, K. & Parkes, L. M. A comparison of
877 dual gradient-echo and spin-echo fMRI of the inferior temporal lobe. *Hum.*
878 *Brain Mapp.* **35**, 4118–28 (2014).
- 879 45. Evans, S. What has replication ever done for us? Insights from neuroimaging
880 of speech perception. *Front. Hum. Neurosci.* **11**, (2017).
- 881 46. Evans, S. & McGettigan, C. Comprehending auditory speech: previous and
882 potential contributions of functional MRI. *Lang. Cogn. Neurosci.* **32**, (2017).
- 883 47. von Kriegstein, K., Eger, E., Kleinschmidt, A. & Giraud, A. L. Modulation of
884 neural responses to speech by directing attention to voices or verbal content.
885 *Cogn. Brain Res.* **17**, 48–55 (2003).
- 886 48. Evans, S. & Davis, M. H. Hierarchical organization of auditory and motor
887 representations in speech perception: Evidence from searchlight similarity
888 analysis. *Cereb. Cortex* **25**, (2015).
- 889 49. Połczyńska, M. M., Japardi, K. & Bookheimer, S. Y. Lateralizing language
890 function with pre-operative functional magnetic resonance imaging in early
891 proficient bilingual patients. *Brain Lang.* **170**, 1–11 (2017).
- 892 50. Emmorey, K., Mehta, S., McCullough, S. & Grabowski, T. J. The neural circuits
893 recruited for the production of signs and fingerspelled words. *Brain Lang.* **160**,
894 30–41 (2016).
- 895 51. Davis, C. J. N-Watch: A program for deriving neighbourhood size and other

- 896 psycholinguistic statistics. *Behav. Res. Methods* **37**, 65–70 (2005).
- 897 52. Kuperman, V., Stadthagen-Gonzalez, H. & Brysbaert, M. Age-of-acquisition
898 ratings for 30,000 English words. *Behav. Res. Methods* **44**, 978–990 (2012).
- 899 53. Wilson, M. The MRC Psycholinguistic Database: Machine Readable
900 Dictionary, Version 2. *Behav. Res. Methods, Instruments Comput.* **20**, 6–11
901 (1988).
- 902 54. Cortese, M. J. & Fugett, A. Imageability ratings for 3 , 000 monosyllabic words.
903 *Behav. Res. Methods, Instruments, Comput.* **36**, 384–387 (2004).
- 904 55. Vinson, D. P., Cormier, K., Denmark, T., Schembri, A. & Vigliocco, G. The
905 British Sign Language (BSL) norms for age of acquisition, familiarity, and
906 iconicity. *Behav. Res. Methods* **40**, 1079–1087 (2008).
- 907 56. Levenshtein, V. I. Binary codes capable of correcting deletions, insertions, and
908 reversals. *Soviet Physics Doklady* **10**, 707–710 (1966).
- 909 57. Fenlon, J., Cormier, K. & Schembri, A. Building BSL SignBank: The lemma
910 dilemma revisited. *Int. J. Lexicogr.* **28**, 169–206 (2015).
- 911 58. Kriegeskorte, N. & Mur, M. Inverse MDS: Inferring dissimilarity structure from
912 multiple item arrangements. *Front. Psychol.* **3**, 1–13 (2012).
- 913 59. Nili, H. *et al.* A toolbox for representational similarity analysis. *PLoS Comput.*
914 *Biol.* **10**, e1003553 (2014).
- 915 60. Walther, A. *et al.* Reliability of dissimilarity measures for multi-voxel pattern
916 analysis. *Neuroimage* (2015). doi:10.1016/j.neuroimage.2015.12.012
- 917 61. Kriegeskorte, N., Goebel, R. & Bandettini, P. Information-based functional

RUNNING HEAD: Influence of modality on conceptual representations

- 918 brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3863–3868 (2006).
- 919 62. Poldrack, R. *et al.* Scanning the Horizon: challenges and solutions for
920 neuroimaging research. *Clim. Chang. 2013 - Phys. Sci. Basis* **1**, 1–30 (2016).
- 921 63. Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F. & Baker, C. I. Circular
922 analysis in systems neuroscience: the dangers of double dipping. *Nat.*
923 *Neurosci.* **12**, 535–540 (2009).
- 924 64. Walker, D. A. JMASM9: Converting Kendall's Tau For Correlational Or Meta-
925 Analytic Analyses. *J. Mod. Appl. Stat. Methods* **2**, 525–530 (2017).
- 926 65. Kaplan, J. T., Man, K. & Greening, S. G. Multivariate cross-classification:
927 applying machine learning techniques to characterize abstraction in neural
928 representations. *Front. Hum. Neurosci.* **9**, 151 (2015).
- 929 66. Nichols, T., Brett, M., Andersson, J., Wager, T. & Poline, J. B. Valid
930 conjunction inference with the minimum statistic. *Neuroimage* **25**, 653–660
931 (2005).

932

933