A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain

- Alexander G. Huth[1],
- Shinji Nishimoto[1],
- An T. Vu[2],
- Jack L. Gallant[1, 2, 3], ,
- [1] Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA 94720, USA
- [2] Program in Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA
- [3] Department of Psychology, University of California, Berkeley, Berkeley, CA 94720, USA

Summary

Humans can see and name thousands of distinct object and action categories, so it is unlikely that each category is represented in a distinct brain area. A more efficient scheme would be to represent categories as locations in a continuous semantic space mapped smoothly across the cortical surface. To search for such a space, we used fMRI to measure human brain activity evoked by natural movies. We then used voxelwise models to examine the cortical representation of 1,705 object and action categories. The first few dimensions of the underlying semantic space were recovered from the fit models by principal components analysis. Projection of the recovered semantic space onto cortical flat maps shows that semantic selectivity is organized into smooth gradients that cover much of visual and nonvisual cortex. Furthermore, both the recovered semantic space and the cortical organization of the space are shared across different individuals.

Video Abstract

Help with MP4 files

Options

Highlights

► The brain represents object and action categories within a continuous semantic space ► This semantic space is organized into broad gradients across the cortical surface ► This semantic space is shared across different individuals

Figures and tables from this article:

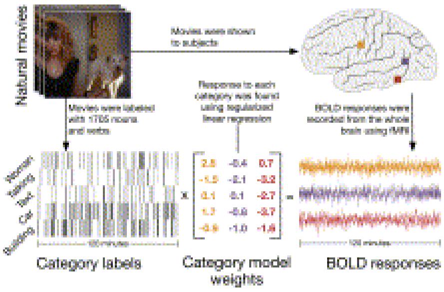

Figure 1. Schematic of the Experiment and ModelSubjects viewed 2 hr of natural movies while BOLD responses were measured using fMRI. Objects and actions in the movies were labeled using 1,364 terms from the WordNet lexicon (Miller, 1995). The hierarchical "is a" relationships defined by WordNet were used to infer the presence of 341 higher-order categories, providing a total of 1,705 distinct category labels. A regularized, linearized finite impulse response regression model was then estimated for each cortical voxel recorded in each subject's brain (Kay et al., 2008; Mitchell et al., 2008;Naselaris et al., 2009; Nishimoto et al., 2011). The resulting category model weights describe how various object and action categories influence BOLD signals recorded in each voxel. Categories with positive weights tend to increase BOLD, while those with negative weights tend to decrease BOLD. The response of a voxel to a particular scene is predicted as the sum of the weights for all categories in that scene.

Figure options

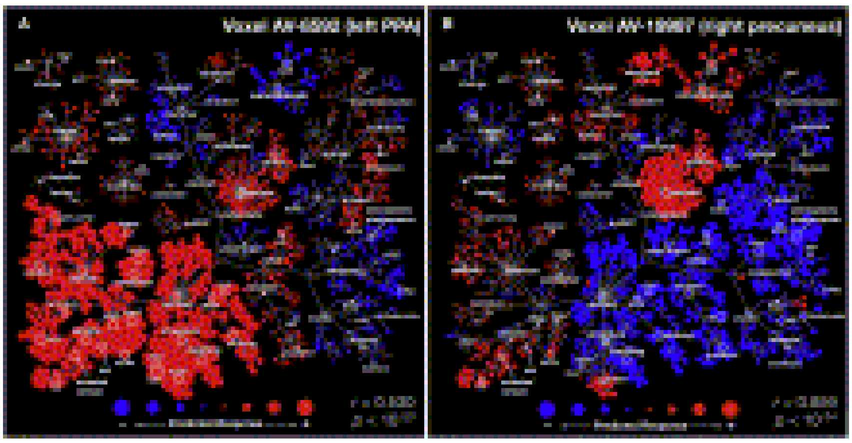

Figure 2. Category Selectivity for Two Individual VoxelsEach panel shows the predicted response of one voxel to each of the 1,705 categories, organized according to the graphical structure of WordNet. Links indicate "is a" relationships (e.g., an athlete is a person); some relationships used in the model are omitted for clarity. Each marker represents a single noun

(circle) or verb (square). Red markers indicate positive predicted responses and blue markers indicate negative predicted responses. The area of each marker indicates predicted response magnitude. The prediction accuracy of each voxel model, computed as the correlation coefficient (*r*) between predicted and actual responses, is shown in the bottom right of each panel along with model significance (see Results for details).(A) Category selectivity for one voxel located in the left hemisphere parahippocampal place area (PPA). The category model predicts that movies will evoke positive responses when “structures,” “buildings,” “roads,” “containers,” “devices,” and “vehicles” are present. Thus, this voxel appears to be selective for scenes that contain man-made objects and structures (Epstein and Kanwisher, 1998).(B) Category selectivity for one voxel located in the right hemisphere precuneus (PrCu). The category model predicts that movies will evoke positive responses from this voxel when “people,” “carnivores,” “communication verbs,” “rooms,” or “vehicles” are present and negative responses when movies contain “atmospheric phenomena,” “locations,” “buildings,” or “roads.” Thus, this voxel appears to be selective for scenes that contain people or animals interacting socially (Iacoboni et al., 2004).

Figure options

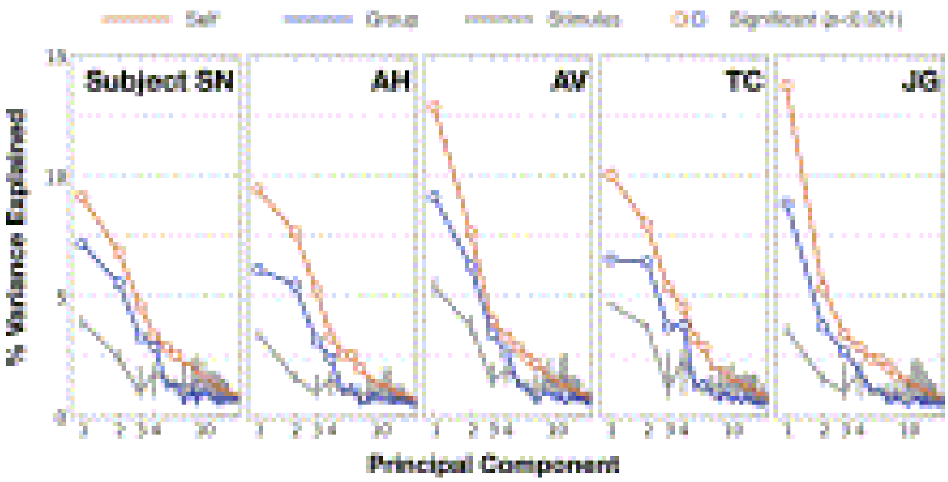

Figure 3. Amount of Model Variance Explained by Individual Subject and Group Semantic SpacesPrincipal components analysis (PCA) was used to recover a semantic space from category model weights in each subject. Here we show the variance explained in the category model weights by each of the 20 most important PCs. Orange lines show the amount of variance explained in category model weights by each subject’s own PCs and blue lines show the variance explained by PCs of combined data from other subjects. Gray lines show the variance explained by the stimulus PCs, which serve as an appropriate null hypothesis (see text and Experimental Procedures for details). Error bars indicate 99% confidence intervals (the confidence intervals for the subjects’ own PCs and group PCs are very small). Hollow markers indicate subject or group PCs that explain significantly more variance ($p < 0.001$, bootstrap test) than the stimulus PCs. The first four group PCs explain significantly more variance than the stimulus PCs for four subjects. Thus, the first four group PCs appear to comprise a semantic space that is common across most individuals and that cannot be explained by stimulus statistics. Furthermore, the first six to nine individual subject PCs explain significantly more variance than the stimulus PCs ($p < 0.001$, bootstrap test). This suggests that while the subjects share broad aspects of semantic representation, finer-scale semantic representations are subject specific.

Figure options

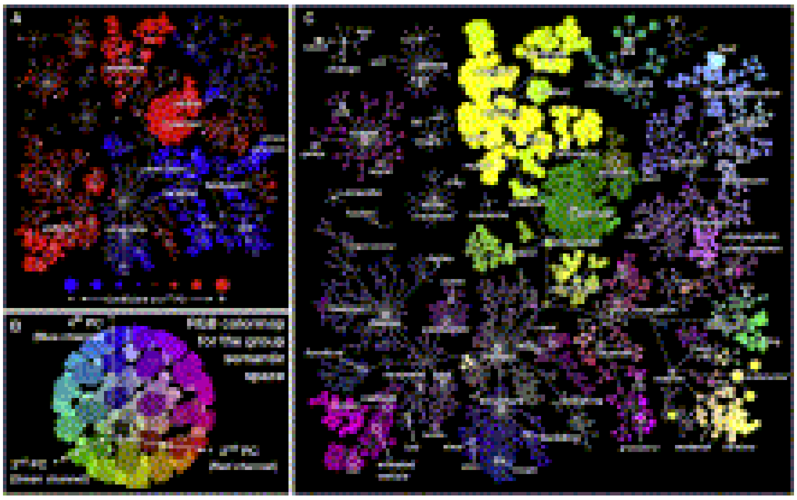

Figure 4. Graphical Visualization of the Group Semantic Space(A) Coefficients of all 1,705 categories in the first group PC, organized according to the graphical structure of WordNet. Links indicate “is a” relationships (e.g., an athlete is a person); some relationships used in the model have been omitted for clarity. Each marker represents a single noun (circle) or verb (square). Red markers indicate positive coefficients and blue indicates negative coefficients. The area of each marker indicates the magnitude of the coefficient. This PC distinguishes between categories with high stimulus energy (e.g., moving objects like “person” and “vehicle”) and those with low stimulus energy (e.g., stationary objects like “sky” and “city”).(B) The three-dimensional RGB colormap used to visualize PCs 2–4. The category coefficient in the second PC determined the value of the red channel, the third PC determined the green channel, and the fourth PC determined the blue channel. Under this scheme, categories that are represented similarly in the brain are assigned similar colors. Categories with zero coefficients appear neutral gray.(C) Coefficients of all 1,705 categories in group PCs 2–4, organized according to the WordNet graph. The color of each marker is determined by the RGB colormap in (B). Marker sizes reflect the magnitude of the three-dimensional coefficient vector for each category. This graph shows that categories thought to be semantically related (e.g., “athletes” and “walking”) are represented similarly in the brain.

Figure options

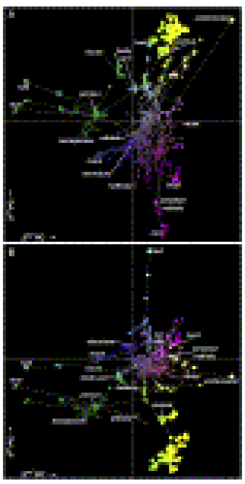

Figure 5. Spatial Visualization of the Group Semantic Space(A) All 1,705 categories, organized by their coefficients on the second and third PCs. Links indicate “is a” relationships (e.g., an athlete is a person) from the WordNet graph; some relationships used in the model have been omitted for clarity. Each marker represents a single noun (circle) or verb (square). The color of each marker is determined by an RGB color map based on the category coefficients in PCs 2–4 (see Figure 4B for details). The position of each marker is also determined by the PC coefficients: position on the x axis is determined by the coefficient on the second PC and position on the y axis is determined by the coefficient on the third PC. This ensures that categories that are represented similarly in the brain appear near each other. The area of each marker indicates the magnitude of the PC coefficients for that category; more important or strongly represented categories have larger coefficients. The categories “man,” “talk,” “text,” “underwater,” and “car” have the largest coefficients on these PCs.(B) All 1,705 categories, organized by their coefficients on the second and fourth PCs. Format is the same as (A). The large group of “animal” categories has large PC coefficients and is mainly distinguished by the fourth PC. Human categories appear to span a continuum. The category “person” is very

close to indoor categories such as "room" on the second and third PCs but different on the fourth. The category "athlete" is close to vehicle categories on the second and third PCs but is also close to "animal" on the fourth PC. These semantically related categories are represented similarly in the brain, supporting the hypothesis of a smooth semantic space. However, these results also show that some categories (e.g., "talk," "man," "text," and "car") appear to be more important than others. Movie S1 shows this semantic space in 3D.

A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain

- Alexander G. Huth[1],
- Shinji Nishimoto[1],
- An T. Vu[2],
- Jack L. Gallant[1, 2, 3],
- [1] Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA 94720, USA
- [2] Program in Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA
- [3] Department of Psychology, University of California, Berkeley, Berkeley, CA 94720, USA

Summary

Humans can see and name thousands of distinct object and action categories, so it is unlikely that each category is represented in a distinct brain area. A more efficient scheme would be to represent categories as locations in a continuous semantic space mapped smoothly across the cortical surface. To search for such a space, we used fMRI to measure human brain activity evoked by natural movies. We then used voxel wise models to examine the cortical representation of 1,705 object and action categories. The first few dimensions of the underlying semantic space were recovered from the fit models by principal components analysis. Projection of the recovered semantic space onto cortical flat maps shows that semantic selectivity is organized into smooth gradients that cover much of visual and nonvisual cortex. Furthermore, both the recovered semantic space and the cortical organization of the space are shared across different individuals.

Video Abstract

Help with MP4 files

Options

Highlights

► The brain represents object and action categories within a continuous semantic space ► This semantic space is organized into broad gradients across the cortical surface ► This semantic space is shared across different individuals

---

Figures and tables from this article:

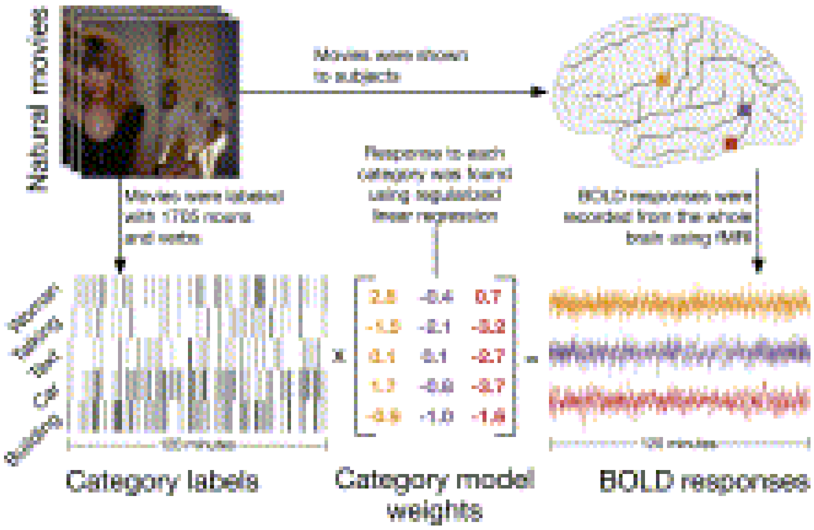

Figure 1. Schematic of the Experiment and ModelSubjects viewed 2 hr of natural movies while BOLD responses were measured using fMRI. Objects and actions in the movies were labeled using 1,364 terms from the WordNet lexicon (Miller, 1995). The hierarchical "is a" relationships defined by WordNet were used to infer the presence of 341 higher-order categories, providing a total of 1,705 distinct category labels. A regularized, linearized finite impulse response regression model was then estimated for each cortical voxel recorded in each subject's brain (Kay et al., 2008; Mitchell et al., 2008;Naselaris et al., 2009; Nishimoto et al., 2011). The resulting category model weights describe how various object and action categories influence BOLD signals recorded in each voxel. Categories with positive weights tend to increase BOLD, while those with negative weights tend to decrease BOLD. The response of a voxel to a particular scene is predicted as the sum of the weights for all categories in that scene.

Figure options

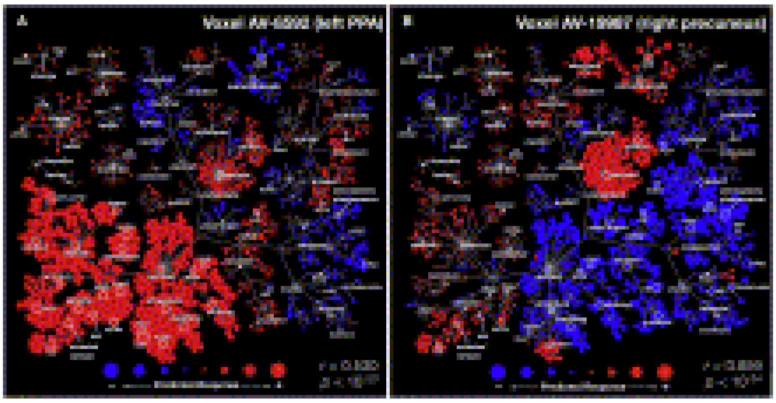

Figure 2. Category Selectivity for Two Individual VoxelsEach panel shows the predicted response of one voxel to each of the 1,705 categories, organized according to the graphical structure of WordNet. Links indicate "is a" relationships (e.g., an athlete is a person); some relationships used in the model are omitted for clarity. Each marker represents a single noun (circle) or verb (square). Red markers indicate positive predicted responses and blue markers indicate negative predicted responses. The area of each marker indicates predicted response magnitude. The prediction accuracy of each voxel model, computed as the correlation coefficient (*r*) between predicted and actual responses, is shown in the bottom right of each panel along with model significance (see Results for details).(A) Category selectivity for one voxel located in the left hemisphere parahippocampal place area (PPA). The category model predicts that movies will evoke positive responses when "structures," "buildings," "roads," "containers," "devices," and "vehicles" are present. Thus, this voxel appears to be selective for scenes that contain man-made objects and structures (Epstein and Kanwisher, 1998).(B) Category selectivity for one voxel located in the right hemisphere precuneus (PrCu). The category model predicts that movies will evoke positive responses from this voxel when "people," "carnivores," "communication verbs," "rooms," or "vehicles" are present and negative responses when movies contain "atmospheric phenomena," "locations," "buildings," or "roads." Thus, this voxel appears to be selective for scenes that contain people or animals interacting socially (Iacoboni et al., 2004).

Figure options

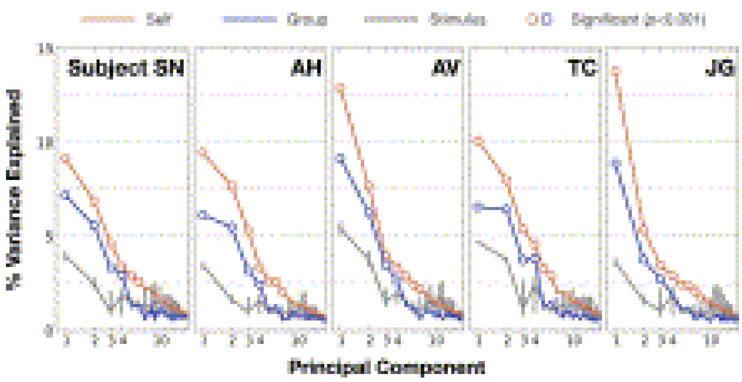

Figure 3. Amount of Model Variance Explained by Individual Subject and Group Semantic SpacesPrincipal components analysis (PCA) was used to recover a semantic space from category model weights in each subject. Here we show the variance explained in the category model weights by each of the 20 most important PCs. Orange lines show the amount of variance explained in category model weights by each subject's own PCs and blue lines show the variance explained by PCs of combined data from other subjects. Gray lines show the variance explained by the stimulus PCs, which serve as an

appropriate null hypothesis (see text and Experimental Procedures for details). Error bars indicate 99% confidence intervals (the confidence intervals for the subjects' own PCs and group PCs are very small). Hollow markers indicate subject or group PCs that explain significantly more variance ($p < 0.001$, bootstrap test) than the stimulus PCs. The first four group PCs explain significantly more variance than the stimulus PCs for four subjects. Thus, the first four group PCs appear to comprise a semantic space that is common across most individuals and that cannot be explained by stimulus statistics. Furthermore, the first six to nine individual subject PCs explain significantly more variance than the stimulus PCs ($p < 0.001$, bootstrap test). This suggests that while the subjects share broad aspects of semantic representation, finer-scale semantic representations are subject specific.

Figure options

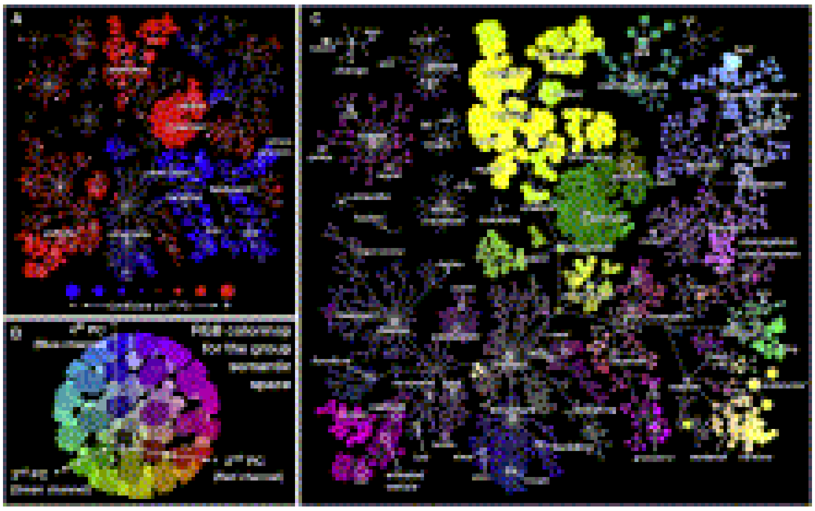

Figure 4. Graphical Visualization of the Group Semantic Space(A) Coefficients of all 1,705 categories in the first group PC, organized according to the graphical structure of WordNet. Links indicate "is a" relationships (e.g., an athlete is a person); some relationships used in the model have been omitted for clarity. Each marker represents a single noun (circle) or verb (square). Red markers indicate positive coefficients and blue indicates negative coefficients. The area of each marker indicates the magnitude of the coefficient. This PC distinguishes between categories with high stimulus energy (e.g., moving objects like "person" and "vehicle") and those with low stimulus energy (e.g., stationary objects like "sky" and "city").(B) The three-dimensional RGB colormap used to visualize PCs 2–4. The category coefficient in the second PC determined the value of the red channel, the third PC determined the green channel, and the fourth PC determined the blue channel. Under this scheme, categories that are represented similarly in the brain are assigned similar colors. Categories with zero coefficients appear neutral gray.(C) Coefficients of all 1,705 categories in group PCs 2–4, organized according to the WordNet graph. The color of each marker is determined by the RGB colormap in (B). Marker sizes reflect the magnitude of the three-dimensional coefficient vector for each category. This graph shows that categories thought to be semantically related (e.g., "athletes" and "walking") are represented similarly in the brain.

Figure options

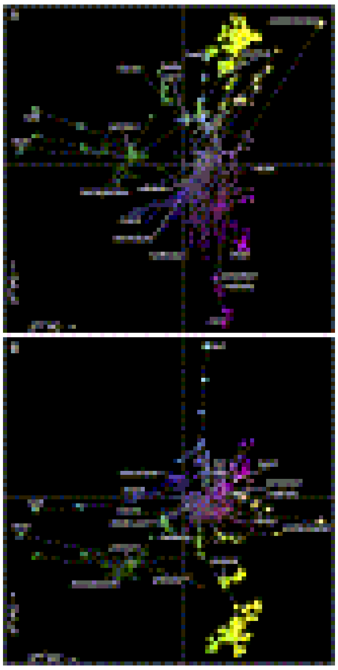

Figure 5. Spatial Visualization of the Group Semantic Space(A) All 1,705 categories, organized by their coefficients on the second and third PCs. Links indicate “is a” relationships (e.g., an athlete is a person) from the WordNet graph; some relationships used in the model have been omitted for clarity. Each marker represents a single noun (circle) or verb (square). The color of each marker is determined by an RGB colormap based on the category coefficients in PCs 2–4 (seeFigure 4B for details). The position of each marker is also determined by the PC coefficients: position on the x axis is determined by the coefficient on the second PC and position on the y axis is determined by the coefficient on the third PC. This ensures that categories that are represented similarly in the brain appear near each other. The area of each marker indicates the magnitude of the PC coefficients for that category; more important or strongly represented categories have larger coefficients. The categories “man,” “talk,” “text,” “underwater,” and “car” have the largest coefficients on these PCs.(B) All 1,705 categories, organized by their coefficients on the second and fourth PCs. Format is the same as (A). The large group of “animal” categories has large PC coefficients and is mainly distinguished by the fourth PC. Human categories appear to span a continuum. The category “person” is very close to indoor categories such as “room” on the second and third PCs but different on the fourth. The category “athlete” is close to vehicle categories on the second and third PCs but is also close to “animal” on the fourth PC. These semantically related categories are represented similarly in the brain, supporting the hypothesis of a smooth semantic space. However, these results also show that some categories (e.g., “talk,” “man,” “text,” and “car”) appear to be more important than others. Movie S1 shows this semantic space in 3D.