



ELSEVIER

Contents lists available at ScienceDirect

Cognition

journal homepage: [www.elsevier.com/locate/cognit](http://www.elsevier.com/locate/cognit)

COGNITION

## Original Articles

## There's more to “sparkle” than meets the eye: Knowledge of vision and light verbs among congenitally blind and sighted individuals

Marina Bedny<sup>a,b,\*</sup>, Jorie Koster-Hale<sup>a</sup>, Giulia Elli<sup>b</sup>, Lindsay Yazzolino<sup>a,b</sup>, Rebecca Saxe<sup>a</sup><sup>a</sup> Massachusetts Institute of Technology, United States<sup>b</sup> Johns Hopkins University, United States

## ARTICLE INFO

## Keywords:

Verb meaning  
Semantic  
Concept  
Blindness  
Experience  
Semantic similarity space

## ABSTRACT

We examined the contribution of first-person sensory experience to concepts by comparing the meanings of perception (visual/tactile) and emission (light/sound) verbs among congenitally blind ( $N = 25$ ) and sighted speakers ( $N = 22$ ). Participants judged semantic similarity for pairs of verbs referring to events of visual (e.g. *to peek*), tactile (e.g. *to feel*) and amodal perception (e.g. *to perceive*) as well as light (e.g. *to shimmer*) and sound (e.g. *to boom*) emission and manner of motion (*to roll*) (total word pairs,  $N = 2041$ ). Relative to the sighted, blind speakers had higher agreement among themselves on touch perception and sound emission verbs. However, for visual verbs, the judgments of blind and sighted participants were indistinguishable, both in the semantic criteria used and subject-wise variability. Blind and sighted individuals alike differentiate visual perception verbs from verbs of touch and amodal perception and differentiate among acts of visual perception e.g. intense/continuous from brief acts of looking (e.g. *peek* vs. *stare*). Light emission verbs are differentiated according to intensity (*blaze* vs. *glow*) and stability (*blaze* vs. *flash*). Thus detailed knowledge of visual word meanings is acquired without first-person sensory access.

## 1. Introduction

In what sense is our notion of a *glowing star* derived from seeing it with our eyes? Do you know what *glowing* is if you have never seen it? Studies with individuals who are blind from birth provide insights into this puzzle. Languages, such as English, have a rich vocabulary for denoting visual referents including color adjectives (e.g. *blue*), verbs of visual perception (e.g. *to peek*) and light emission events (e.g. *to sparkle*) (Winter, Perlman, & Majid, 2018). What are the meanings of these terms, for people who have never seen?

British empiricist philosophers engaged in thought experiments about blindness to test intuitions about the origins of knowledge and concluded that blind people and sighted people must have very different concepts (1732; Berkeley, 1948; Hobbes, 1641, 1984; Hume, 1739, 1978; Locke and Nidditch, 2011). Following in their footsteps, early educational psychologists dubbed blind individuals' use of words for visual categories such as colors and light events “verbalisms,” because of the words' alleged meaninglessness (Cutsforth, 1932, 1951, see Rosel, Caballer, Jara, & Oliver, 2005 for review of verbalism literature). Contrary to such ideas, Landau and Gleitman (1985) showed that blind preschoolers can use color adjectives and visual perception verbs in appropriate ways – both when referring to themselves and when

referring to sighted people. For example, Kelly, a blind four-year-old, responded to the instruction of *look* by holding out her hands. When asked to make it so her mother couldn't see an object, Kelly hid it in her pocket. Kelly also understood that colors were physical properties that sighted people could perceive, but she could not. These results suggested that from an early age, blind children can meaningfully comprehend and produce visual terms.

Nevertheless, the question remains: just how rich is blind individuals' knowledge about vision and how similar is it to the knowledge of sighted people? A challenge in answering these questions is how to measure, and quantitatively compare, the concepts of blind and sighted individuals. One possibility is to ask for an explicit definition of the words. Landau and Gleitman (1985) asked a congenitally blind adult to provide definitions of twenty verbs related to visual experience. Her definitions showed both appropriate knowledge of the meanings of the words and sensitivity to their use in visual contexts. For example, she defined *to notice* as: “to see something that comes into your view. But not only to see it, but to perceive it and understand it. You could sit on this rocking chair and not notice the color of it at all.” Similarly, Lenci, Baroni, Cazzolli, and Marotta (2013) recently collected feature norms for 5 verbs of visual perception (in Italian, glossed as *spot*, *glimpse*, *peep*, *catch sight of*, and *peer at*) from congenitally blind and

\* Corresponding author at: Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, United States.  
E-mail address: [Marina.bedny@jhu.edu](mailto:Marina.bedny@jhu.edu) (M. Bedny).

<https://doi.org/10.1016/j.cognition.2019.03.017>

Received 19 December 2018; Received in revised form 22 March 2019; Accepted 24 March 2019

Available online 30 March 2019

0010-0277/ © 2019 Elsevier B.V. All rights reserved.

sighted native Italian speakers. Again, the blind individuals generated reasonable meaning-features for all of these verbs. For instance, *to peep* was associated with the features “to watch,” “something,” “secretly,” and “not to be seen”, whereas *to spot* was associated with the features “to see”, “something”, “far away” and “distance” by both sighted and blind people. The disadvantage of free responses, however, is that these data are very sparse, as features may be rarely mentioned, or described using homonyms. Thus, it remains hard to quantitatively test whether blind individuals’ meanings of *to peep* or *to spot* are different from those of sighted individuals.

More generally, people may have very rich and detailed knowledge of the meanings of words, but not reveal that knowledge in their definitions. The pragmatics of the task may cause people to limit the features they generate to relatively distinctive properties within an implied context. For example, people may be more likely to volunteer that zebras have stripes than that they have mouths. Shepard and Chipman (1970) argued that people “seem unable to tell us anything significant about the structure of an individual mental [representation] as such. What they can, however, tell us about is the relations between that internal representation and other internal representations.” That is a practical, albeit incomplete, way to elicit rich information about the meaning of a word is to ask people to produce not the meaning itself but judgments of how it relates to the meanings of other words. “Thus, we easily report that orange is more similar to red than to blue without being able to say anything significant [...] about the unique subjective experience of the color orange itself” (Shepard & Chipman, 1970).

Estimates of the similarity between pairs of word-meanings are easy to elicit from a wide range of domains and are naturally quantitative. People make highly stable, reliable judgments of the similarities of pairs of mammals (Rumelhart & Abrahamson, 1973), birds (Rips, Shoben, & Smith, 1973), fruits (Hutchinson & Lockhead, 1977), foods (Ross & Murphy, 1999), numbers (Shepard, Kilpatrick, & Cunningham, 1975), colors (Shepard & Cooper, 1992), emotions (Roberts & Wedell, 1994), and personality types (Bimler & Kirkland, 2007), among many other examples. The semantic similarity spaces derived from such judgments do not provide a complete measure of what people know about a domain. For example, the similarity space of animals may not reflect people’s knowledge about their diets, or how the animals are used by humans for making food or clothing, unless people are asked explicitly to judge these particular features (e.g. Medin et al., 2002, Tenenbaum & Griffiths, 2001, Murphy, 2004). Nevertheless, semantic similarity judgments capture a large amount of information quickly and quantitatively and predict performance on more implicit tasks for the same words, such as memory confusions and priming effects, suggesting that similarity judgments capture some stable semantic properties of words (e.g. Hutchinson & Lockhead, 1977).

Critical for the present purposes, semantic similarity judgments are sensitive to between-group differences in semantic knowledge. For example, changes in similarity judgments provide an early signal of cognitive deterioration in patients with Alzheimer’s Disease (AD). Disruption of similarity judgments predicts the rate of patients’ loss of cognitive function over the following year (Chan et al., 1993). Semantic similarity judgments are also sensitive to changes in culture and experience. For instance, the perceived similarity of mammals reveals both substantial agreement, and significant deviations, between American college students and Itzaj Mayan adults (Lopez, Atran, Coley, Medin, & Smith, 1997).

Thus, semantic similarity ratings of pairs of words, while only a partial measure of what people know, offer a way to quantitatively compare the meanings of vision-related words among blind and sighted individuals. Indeed, this logic has been used by previous studies to test blind individuals’ knowledge of color. On average both sighted and blind groups show a systematic pattern, resembling a color wheel: red is similar to orange which is similar to yellow, and so on, until violet, which is similar to blue and red (Shepard & Cooper, 1992). There is more variation across blind than sighted adults, however, some blind

adults reproduce the color wheel, while others make idiosyncratic judgments with large deviations from the typical pattern (Marmor, 1978; Saysani, Corballis, & Corballis, 2018; Shepard & Cooper, 1992). These results suggest (i) that it is possible to acquire typical knowledge of color similarity without direct first-person experience, but also (ii) that first-person experience is a particularly efficient way of doing so – at least for color.

Given these mixed prior results, it is an open question how generally, and how profoundly, blind individuals’ knowledge of visual words differs from that of sighted people. In particular, it is uncertain whether blind and sighted individuals share detailed knowledge of visual verb meanings. To address this question, we acquired the largest sample to date of similarity ratings for visual verbs from congenitally blind and sighted English speaking adults (see <https://osf.io/zx3t9/> for data). Participants judged the semantic similarity of visual verbs including verbs of visual perception (e.g. *to peek*, *to peer*) and light emission (e.g. *to sparkle*, *to shine*). We chose fifteen verbs from each category, thus including nearly all frequently used visual verbs in the English language (Levin, 1993). English has a fairly large vocabulary of such words relative to other languages (Majid et al., 2018; Winter et al., 2018). Knowledge of visual perception verbs was compared to knowledge of tactile perception (e.g. *to touch*, *to feel*) and amodal knowledge acquisition (e.g. *to perceive*, *to examine*, *to discover*). Light emission verbs were compared to verbs of sound emission, both non-agentive (e.g. *to boom*, *to clank*) and agentive (e.g. *to grunt*, *to shout*). In total, each blind and sighted control participant made 2041 judgments. In addition, we collected a second sample of similarity judgments from workers on Amazon Mechanical Turk. This second sample of sighted data enabled us to get a benchmark of lexical variability across sighted participants. We reasoned that judgments would differ across people due to measurement noise as well as blindness-unrelated individual differences (e.g. education, memory capacity). If blindness systematically affects knowledge of visual verb meanings, then the semantic similarity judgments of a sample of blind individuals should differ more from a sighted sample than two randomly sampled groups of sighted speakers do from each other. If so, sensory experience may have special effects on the lexicon, apart from other individual variation. In sum, the data enable us to measure how first-person sensory experience influences the meanings of words whose referents are sensory.

## 2. Methods

### 2.1. Participants

Twenty-five congenitally blind (20 female) and twenty-two sighted (11 female) participants took part in the experiment. All participants went through a detailed screening interview over the phone and reported having no cognitive or neurological disabilities and being English native speakers (learned English before age 5). Blind participants were totally blind from birth (had at most minimal light perception) and had lost their vision due to abnormalities of the eyes or the optic nerve (not due to brain damage) (Table 1). Sighted and blind participants were matched to each other in age (blind:  $M = 44.86$ ,  $SD = 14$ , missing age information for 3 participants; sighted:  $M = 50.64$ ,  $SD = 8.51$ ) and level of education (blind: ranging from some college (no degree) to Doctoral Degree, Mode = Master’s Degree; sighted: ranging from High School Diploma to Doctoral Degree, Mode = Bachelor’s Degree). Three participants did not provide similarity judgments for one whole semantic category. Thus, we obtained similarity judgments on the perception verbs by 22 sighted and 24 blind participants, and on the emission and manner of motion verbs by 21 sighted and 25 blind participants.

In addition, we obtained data on Amazon Mechanical Turk (AMTurk) from a sighted reference group ( $N = 303$ , henceforth sighted reference group) that was then compared to the ratings of the blind adults and sighted controls. Mechanical Turk participants were all

**Table 1**  
Participants demographic information.

Participant	Gender	Age	Cause of blindness	Light perception	Education level
CB_06	F	25	Retinopathy of Prematurity	None	BA in progress
CB_08	M	42	Retinopathy of Prematurity	Minimal	MA
CB_09	F	47	Retinopathy of Prematurity	None	BA in progress
CB_14	F	63	Retinopathy of Prematurity	Minimal	MA
CB_15	F	41	Retinopathy of Prematurity	Minimal	BA
CB_18	F	59	Hereditary eye condition (unspecified)	None	BA in progress
CB_19	M	35	Norrie's Syndrome	None	BA
CB_22	F	57	Retinopathy of Prematurity	None	BA in progress
CB_23	F	50	Retinopathy of Prematurity	None	MA
CB_24	F	64	Retinopathy of Prematurity	None	MA
CB_25	F	53	Congenital deformation of rods and cones	Minimal	MA
CB_26	F	35	Retinopathy of Prematurity	Minimal	BA
CB_27	F	46	Retinopathy of Prematurity	Minimal	PhD
CB_28	M	49	Born with underdeveloped eyes (unspecified)	None	BA in progress
CB_29	M	–	Retinopathy of Prematurity	None	AA
CB_30	F	44	Microphthalmia	Minimal	MA
CB_31	F	60	Retinopathy of Prematurity	None	BA
CB_32	F	61	Retinoblastomas	None	MA
CB_33	F	27	Retinopathy of Prematurity	None	AA
CB_34	F	–	Retinopathy of Prematurity	None	BA
CB_35	F	27	Retinopathy of Prematurity	None	AA
CB_37	F	58	Retinopathy of Prematurity	None	Professional Degree
CB_38	M	–	Retinopathy of Prematurity	None	MA
CB_40	F	21	Retinopathy of Prematurity	Minimal	BA in progress
CB_41	F	23	Leber's Congenital Amaurosis	Minimal	BA
<b>Average</b>					
Blind (N = 25)	20F	44.86	–	–	MA
Sighted (N = 22)	11F	50.64	–	–	BA

English native speakers from the United States, according to self-report and AMTurk data. No other demographic data were available for these participants. We excluded 37 participants because they either gave the same response to all items, or answered the survey in < 4 min, leaving 266 participants in the analyses. No further demographic information was collected from the participants.

## 2.2. Stimuli

The stimuli consisted of three broad categories of verbs (Table 2). Verb frequencies were obtained from the SubtlexUS database (Brysbaert & New, 2009). We selected verbs from the Levin (1993) text, choosing those that were frequent and likely to be familiar to most speakers. The first category included verbs referring to agentive experiences that were either visual (e.g. *to glance*, *to stare*,  $N = 15$ ), tactile (e.g. *to touch*, *to feel*,  $N = 15$ ) or amodal (e.g. *to investigate*, *to notice*,  $N = 15$ ) ( $\log_{10}(\text{freq.})$  visual:  $M = 2.77$ ,  $SD = 1.22$ ; tactile:  $M = 2.65$ ,  $SD = 0.8$ ; amodal:  $M = 2.79$ ,  $SD = 0.9$ ). We henceforth refer to this class as the perception verbs since all of them involved acts of knowledge acquisition. One visual perception verb, *to ogle*, was reported as unfamiliar by most of both sighted and blind participants, and so was excluded from all analyses, leaving 44 perception verbs (*ogle* not included in frequency calculations above). The second class consisted of verbs that refer to events in the environment that are perceptible either through vision only (i.e. light emission, e.g. *to sparkle*, *to shine*,  $N = 15$ ) or hearing only (sound emission e.g. *to buzz*, *to bang*,  $N = 30$ ). Among the sound emission verbs, half referred to sounds generated by animate agents (e.g. *to bark*) and half by inanimate objects (e.g. *to clang*) ( $\log_{10}(\text{freq.})$  light:  $M = 2.02$ ,  $SD = 0.54$ ; animate sound:  $M = 1.82$ ,  $SD = 0.56$ ; inanimate sound:  $M = 1.88$ ,  $SD = 0.65$ ). The third category included manner of motion verbs (e.g. *to hobble*, *to roll*,  $N = 15$ ;  $\log_{10}(\text{freq.})$   $M = 2.11$ ,  $SD = 0.78$ ). In addition to the above described verbs, participants also judged 15 mental verbs (e.g. *to enjoy*, *to tolerate*) but these were not relevant to the hypotheses of the current study and were thus not included in the reported analyses. All verbs were presented as infinitives (i.e. preceded by *to*). Note that although no attempt was made to include all possible verbs, the stimuli include most of the

frequent visual verbs within the English language.

Verb-pairs were constructed by making all possible pairings within the broad semantic categories. These pairings avoided putting together verbs that were highly dissimilar in meaning and thus required participants to make more fine-grained judgments. Perception verbs were paired with each other within and across modalities (*to peek* – *to stare*, *to touch* – *to see*, total 946 pairs). Among verbs that described perceptible events, all emission verbs were paired with each other both within modality (e.g. *to buzz* – *to ring*) and across modalities, (*to buzz* – *to sparkle*, total 990 pairs). Motion verbs were paired only amongst themselves (105 pairs.)

## 2.3. Task

Blind and sighted participants completed an online survey in which they rated the semantic similarity of verb pairs (e.g. *to see* – *to touch*) on a scale from 1 (not at all similar) to 7 (very similar). Blind participants listened to the stimuli using commercially available screen-readers; sighted participants read written words on a computer screen. Each participant was asked to rate all possible pairings of verbs within each broad semantic class for a total of 2041 word pairs per participant (including pairs with *to ogle* and the 15 mental verbs, which were later dropped from analysis). The survey took a total of 6–8 h. Participants completed the survey from home over the span of 2–4 weeks. Each participant was given a user name and password. They would log into their account, complete a portion of the survey, and then return to it at their convenience.

All participants were walked through the survey instructions either over the phone (blind participants) or in person (sighted controls) (see Supplementary Information for instructions). Participants were asked to use the full scale of 1–7. They were told there were no correct answers and reminded to rate how similar the verbs were in meaning, not in sound or spelling. Before starting the main survey, participants completed a practice session with 100 pairs of animal nouns (e.g. *the bear* – *the tiger*).

The survey was divided into three sections: perceptual experience verbs, emission verbs, and manner of motion verbs. The order of the

**Table 2**  
Complete list of verb stimuli by semantic subcategory. Note that all verbs were preceded by the infinitive marker *to* in the actual experiment.

Perception			
Visual		Touch	Amodal
gawk		caress	characterize
gaze		dab	classify
glance		feel	discover
glimpse		grip	examine
leer		nudge	identify
look		pat	investigate
peek		pet	learn
peer		pinch	note
scan		prod	notice
see		pub	perceive
spot		scrape	question
stare		stroke	recognize
view		tap	scrutinize
watch		tickle	search
ogle*		touch	study

Emission			Manner of Motion
Light	Animate Sound	Inanimate Sound	
blaze	bark	beep	bounce
blink	bellow	boom	float
flare	groan	buzz	glide
flash	growl	chime	hobble
flicker	grumble	clang	roll
gleam	grunt	clank	saunter
glimmer	howl	click	scurry
glint	moan	crackle	skip
glisten	mutter	creak	slither
glitter	shout	crunch	spin
glow	squawk	gurgle	strut
shimmer	wail	hiss	trot
shine	whimper	sizzle	twirl
sparkle	whisper	squeak	twist
twinkle	yelp	twang	waddle

\* Excluded from analyses as too unfamiliar to participants.

sections was counterbalanced across participants, as was the order of words within a pair (i.e. half of the participants were presented with *to look* – *to stare* and the other half *to stare* – *to look*). The order of verb pairs within each section was randomized across participants. We included catch trials to ensure participants were attending to the task. Vegetable names disguised as verbs (e.g. *to carrot* – *to potato*) appeared 10% of the time. Participants were instructed to enter V for these pairs, rather than a similarity rating.

Participants in the sighted reference group rated subsets of 100–200 word pairs. Each pair was rated by 20 participants. The order of pairs and words was counterbalanced across participants. The AMTurk ratings were combined to generate a single complete dataset and were never analyzed at the single-subject level, since no participant generated a full data set.

## 2.4. Analyses

### 2.4.1. Generating semantic dissimilarity matrices

We created semantic dissimilarity matrices both on the participants' raw similarity judgments and on the normalized scores. The normalized scores were obtained by first z-scoring ( $M = 0$ ,  $SD = 1$ ) the similarity judgments within participants to account for individual differences in Likert scale use, then normalizing  $X_{\text{new}} = (X - X_{\text{min}})/(X_{\text{max}} - X_{\text{min}})$  within the three main semantic categories (perception, emission, manner of motion) such that each verb pair had a similarity distance within [0,1] range.

We generated individual subjects' similarity matrices separately for the three main semantic categories (i.e. perception, emission and manner of motion verbs), as well as for each semantic subcategory (i.e.

sight, touch and amodal perception; light and animate/inanimate sound emission). Group similarity matrices were then created by averaging individual subject matrices across participants for each pair of verbs. Dissimilarity matrices were finally obtained as the maximum value of the scale (7 for the raw data, 1 for the normalized score) minus the similarity matrices (Fig. 1). Data are publicly available at: <https://osf.io/>

### 2.4.2. Across and within-Group agreement

We used blind and sighted participants' normalized similarity ratings to measure within and across group coherence. To quantify the similarity of lexical knowledge between blind and sighted participants, we first computed the Spearman's rho rank correlation between the sighted reference group matrix and the average blind and sighted group dissimilarity matrices, respectively. As the group average matrices are less noisy than the single-subject ones, group-level correlations provide the best between-group similarity estimates. However, they do not take into account within-group variance. To test whether the between-group correlations are reliable across individuals, we correlated (Spearman's rho rank correlation) the sighted reference group matrices with the blind and sighted individual single-subject matrices. We report the Fisher-Z transformed average correlations of each subject to the sighted reference, which gives a measure of how correlated each individual blind or sighted test subject is to a randomly sampled group of sighted subjects. The significance of these correlations was then tested using Student's t-tests across the Fisher-Z transformed single-subject correlations. Correlations were computed on the normalized dissimilarity matrices using the Hmisc package in R (Harrell & Dupont, 2014).

Finally, we measured within-group coherence to compare the

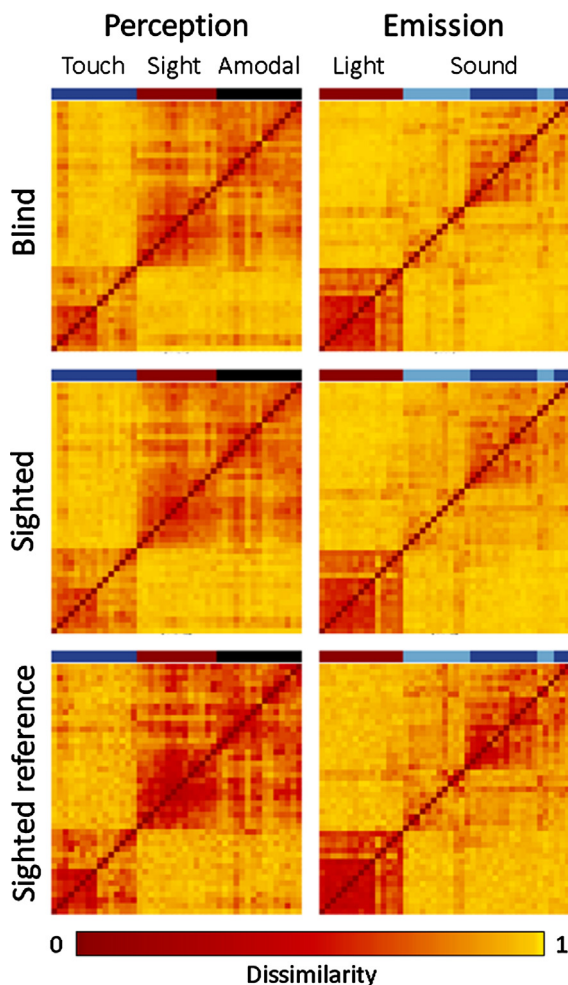


Fig. 1. Group average semantic dissimilarity matrices for perception and emission verbs, blind, sighted and sighted reference groups (top to bottom). Darker red corresponds to more similar.

degree to which blind and sighted speakers agree amongst themselves on the meanings of the tested verbs. First we calculated the Kendall's W Coefficient of Concordance for each verb type and group. The Kendall's W is an estimate of the correlation between all pairs of participants within a group and has previously been used to measure subject agreement on semantic similarity measures (Barsalou & Sewell, 1984; Barsalou, 1987, 1993). Since Kendall's W is a group-wise-metric with no variance, we used a leave-one-subject-out procedure to test for differences in within-group coherence across groups. This analysis correlated responses of each participant to their own group holding their own data out (n-1). The Kendall W and leave-one-subject out procedure produced nearly identical measures of within-group coherence.

#### 2.4.3. Multidimensional scaling

Multidimensional scaling (MDS) was used to visualize the natural clustering of the group dissimilarity matrices using the basic MDS approach as implemented in the SMACOF package in R (Mair, de Leeuw, & Groenen, 2015). Specifically, we computed two-way interval MDS models using the Stress Majorization of a Complicated Function (SMACOF) approach, which minimizes the stress-function by means of an iterative majorization process (i.e. the SMACOF algorithm; De Leeuw & Heiser, 1977, De Leeuw & Mair, 2011). We determined the dimensionality of the models by creating scree plots, which show the stress values as a function of dimensions (ranging between 2 and 10, see Supplementary Fig. 1), and selecting the value at which adding dimensions no longer improves the model fit substantially (i.e. the elbow

of the curve). Based on this heuristic, we report the results of basic MDS models fit with 4 dimensions. For each MDS, the goodness-of-fit is given as the Kruskal's normalized stress-1 value. The smaller the stress value, the better the fit of the solution. The fit significance is evaluated with a permutation test which provides a null distribution of stress values based on the random permutation ( $N = 1000$ ) of the dissimilarity matrices (Mair, de Leeuw & Groenen, 2015).

#### 2.4.4. Hierarchical agglomerative clustering and dendrograms

Data were hierarchically clustered beginning with each observation as its own, single-item cluster and progressively merging clusters up the hierarchy (Murtagh & Legendre, 2014). We used Ward's minimum variance criterion to create clusters, which merges clusters minimizing the total within-cluster variance (i.e. the weighted squared distance between cluster centers) relative to other possible merges (Murtagh & Legendre, 2014). Thus, the dendrogram is built from the bottom up: each verb is originally assigned to its own cluster and, at each step, the two closest clusters (i.e. those whose merged result has the least variance) are merged into a new, larger cluster, eventually converging at the origin of the branching. Distances between cluster centers were recomputed by the Lance-Williams dissimilarity update formula (Murtagh & Legendre, 2014). This procedure tends to lead to compact and spherical clusters. Analyses were done using the pvclust package in R (Suzuki & Shimodaira, 2006), which assesses the clusters' reliability using multiscale bootstrap resampling. This procedure computes several clusters by resampling over specific verb pairs. Then, the reliability across resampling is calculated for each branching and used to generate the final dendrogram. This algorithm is a computationally fast way of implementing standard double-bootstrap, in which the standard error of each bootstrap replication is estimated using bootstrap resampling within the resampled replication.

### 3. Results

#### 3.1. Blind individuals distinguish visual verbs from verbs in other modalities and amodal verbs

For sighted and blind participants, group-wise MDS revealed analogous semantic structures across groups. Fig. 2 shows the first 2 dimensions that emerge for the blind and sighted groups. Perception verbs separate into three major clusters by modality (sight, touch, amodal). Among these clusters, visual verbs (e.g. *to peek*) and amodal verbs (e.g. *to investigate*) are closer (more similar) to each other than to touch verbs (e.g. *to feel*). Likewise, the emission verbs separate according to light versus sound verbs and, among sound verbs, into agentive and non-agentive verbs (Fig. 2). Similar results were obtained when raw similarity scores were compared across groups using standard parametric statistics (see Supplementary Information and Supplementary Fig. 2 for details). For both blind and sighted groups, the basic MDS fits were good, i.e. stress measures were low and comparable across groups for all verb categories (perception verbs stress: sighted 0.12, blind 0.12; emission verbs stress: sighted 0.14, blind 0.13; within-group goodness-of-fit significances  $p$ 's < 0.0001).

#### 3.2. Preserved semantic similarity structure of visual verbs in blind people

##### 3.2.1. Agreement of semantic similarity ratings within and across groups.

Do blind individuals make similar distinctions among visual verbs as sighted individuals? To address this question, we first asked whether the semantic similarity ratings of blind adults were as correlated with those of a group of sighted participants as two independent groups of sighted participants are to each other.

At the group level, the average ratings of blind individuals for visual verbs were as highly correlated with those of the sighted reference group as the two sighted groups were to each other (visual perception verbs: sighted to sighted reference group  $\rho(89) = 0.84$ , blind to

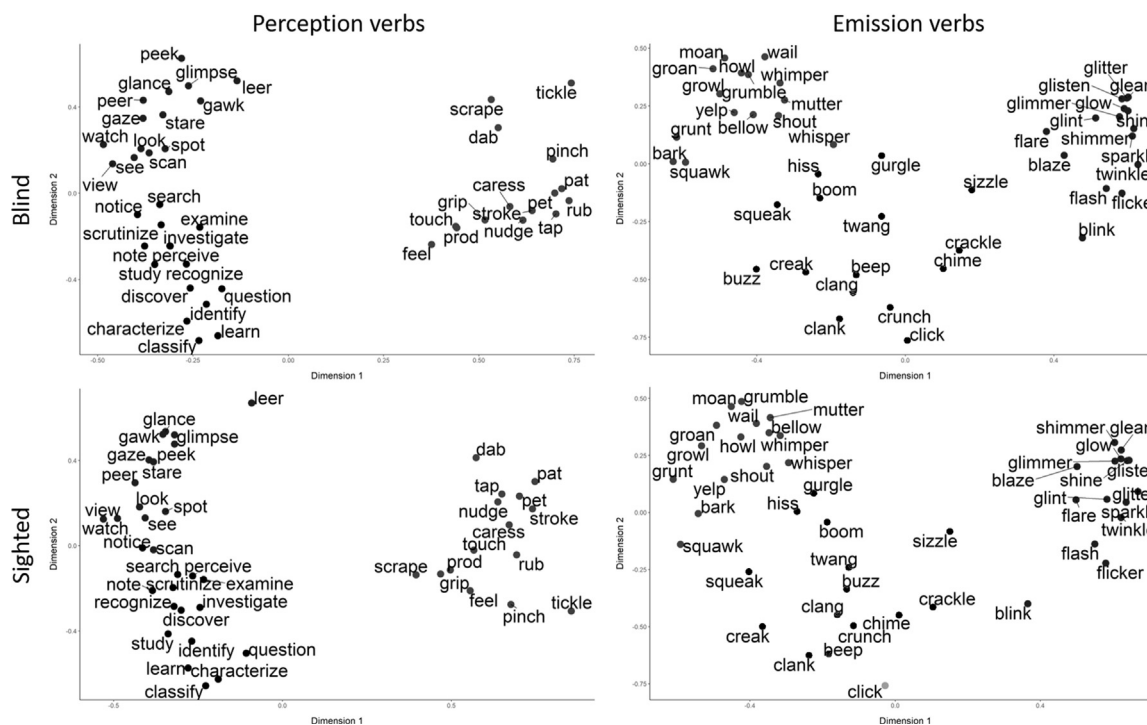


Fig. 2. Blind (top) and sighted (bottom) group MDS results for perception (left) and emission (right) verbs. First two dimensions shown.

sighted reference group  $\rho(89) = 0.81$ ; light emission verbs: sighted to sighted reference group  $\rho(103) = 0.91$ , blind to sighted reference group  $\rho(103) = 0.93$ , Fig. 3A). To take variability across individuals into account, we iteratively correlated the ratings of individual blind and sighted participants to the mean ratings of the sighted reference group. For visual perception verbs, the ratings of blind and sighted individuals were equally well-correlated with those of the sighted reference group (sighted to sighted reference  $M$  Fisher- $z = 0.51$ ,  $SD = 0.12$ ; blind to sighted reference  $M$  Fisher- $z = 0.55$ ,  $SD = 0.15$ ; two-sample  $t$ -test across groups:  $t(41.55) = -0.84$ ,  $p = 0.4$ ).

We observed the same pattern for verbs of light emission. The individual ratings from blind and sighted participants were equally well correlated with those of the sighted reference group (sighted to sighted reference  $M$  Fisher- $z = 0.68$ ,  $SD = 0.16$ ; blind to sighted reference, Mean Fisher- $z = 0.7$ ,  $SD = 0.17$ ; two-sample  $t$ -test across groups:  $t(43.82) = 0.25$ ,  $p = 0.8$ ).

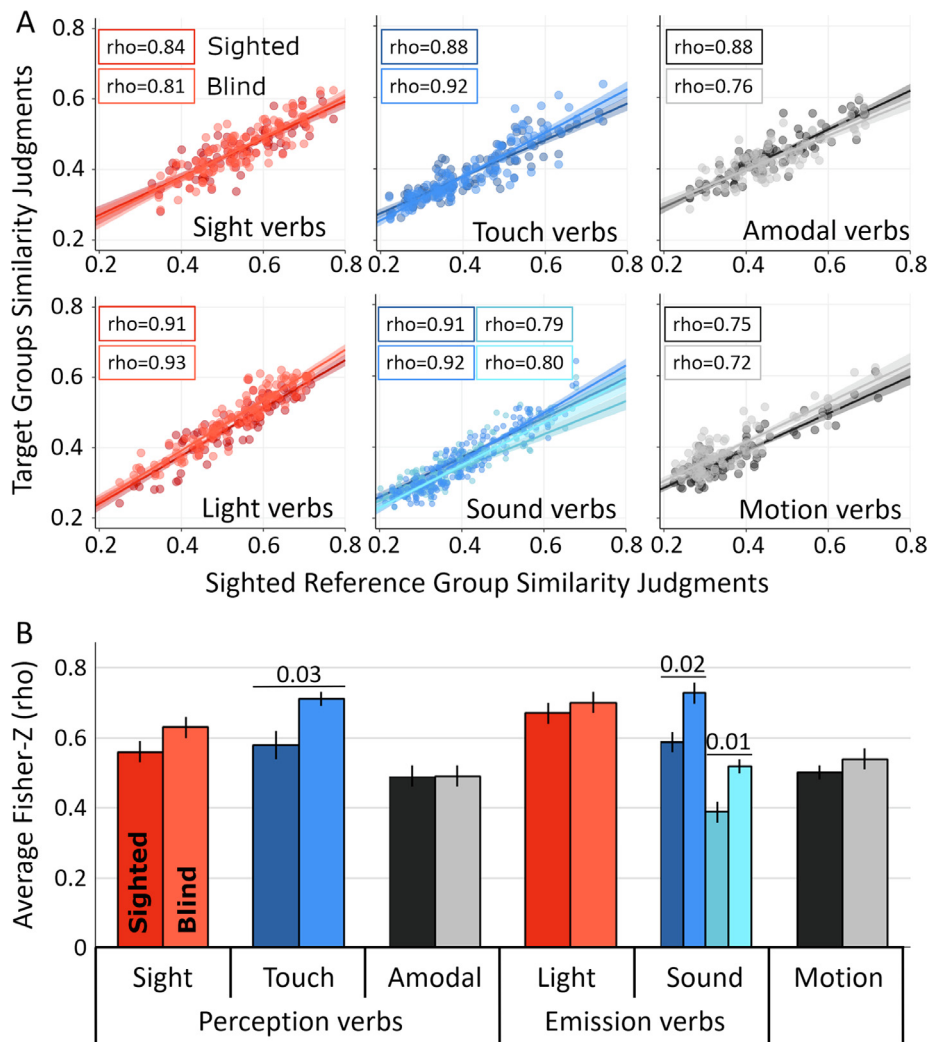
We next asked whether visual verbs were less similar between blind and sighted participants, as compared to amodal verbs. On the contrary, relative to the sighted test group, the ratings of blind participants were slightly, but not significantly more similar to those of the sighted reference group for visual verbs as compared to amodal verbs (2 Group (blind, sighted)  $\times$  2 Modality (visual, amodal) repeated measures ANOVAs on blind to sighted reference and sighted to sighted reference; visual perception vs. amodal perception verbs: main effect of Group  $F(1,44) = 0$ ,  $p = 0.9$ , main effect of Modality  $F(1,44) = 5.39$ ,  $p = 0.25$ , Group  $\times$  Modality interaction  $F(1,44) = 3.14$ ,  $p = 0.08$ ; light emission vs. motion verbs: main effect of Group  $F(1,43) = 0.03$ ,  $p = 0.88$ , main effect of Modality  $F(1,43) = 22.89$ ,  $p < 0.0001$ , Group  $\times$  Modality interaction  $F(1,43) = 0.06$ ,  $p = 0.8$ ).

Finally, we conducted a group coherence analysis to determine whether blind and sighted participants differed with regard to agreement within their own group. First, we computed the Kendall's  $W$  Concordance Coefficient, which is an estimate of the agreement/correlation between every pair of subjects within a group, for each verb-type and subject group (Supplementary Fig. 3, Supplementary Table). Blind and sighted speakers were equally coherent in their judgments for visual verbs (visual perception verbs: sighted  $W = 0.31$ , blind

$W = 0.33$ ; light emission verbs: sighted  $W = 0.38$ , blind  $W = 0.39$ , all  $W$ 's significantly greater than zero,  $p$ 's  $< 0.0001$ ). Because Kendall's  $W$  is a group-wise measure, we used a leave-one-participant-out procedure to estimate variability across participants, correlating each blind and sighted participant to their own group, holding their data out (Fig. 3B). These analyses revealed that blind and sighted participants were equally coherent amongst themselves for verbs of visual perception and light emission (two-sample  $t$ -test blind vs. sighted visual perception verbs  $t(42.86) = 1.3$ ,  $p = 0.2$ , light emission verbs  $t(43.87) = 0.48$ ,  $p = 0.63$ ; visual vs. amodal perception. Blind participants were also no less coherent for visual verbs than for amodal verbs, relative to the sighted either for perception or emission verbs (perception verbs: Group (blind, sighted)  $\times$  Verb Modality (visual, amodal) repeated measures ANOVAs main effect of Group  $F(1,44) = 0.69$ ,  $p = 0.4$ , main effect of Modality  $F(1,44) = 14.72$ ,  $p < 0.0005$ , Group  $\times$  Modality interaction  $F(1,44) = 1.81$ ,  $p = 0.18$ ; emission verbs (light emission vs. motion verbs: main effect of Group  $F(1,43) = 0.74$ ,  $p = 0.4$ , main effect of Modality  $F(1,43) = 18.51$ ,  $p < 0.0005$ , Group  $\times$  Modality interaction  $F(1,43) = 0$ ,  $p = 0.97$ ). Rather, both sighted and blind groups showed higher coherence for the "visual" categories, hence the significant (Fig. 3). These results indicate that the within-category similarity structure of visual verbs is preserved among blind individuals. Two random samples of sighted participants agrees to the same extent as blind and sighted samples. Furthermore, blind individuals show the same degree of within-group coherence for visual groups as sighted people.

### 3.2.2. What do blind individuals know about visual verbs?: MDS and hierarchical clustering analysis

We used MDS as well as hierarchical clustering analyses within the visual verb classes to gain insight into the content of blind and sighted people's knowledge about these words. For sight perception verbs, MDS produced good and comparable fits across groups but did not yield interpretable dimensions (MDS stress value sighted 0.08, blind 0.07, goodness-of-fit each group  $p < 0.0001$ ). Hierarchical clustering analyses revealed that for both blind and sighted groups, visual perception verbs clustered into intense, prolonged acts of seeing (e.g. *to leer*, *to*



**Fig. 3.** A. Correlations of blind (light colors) and sighted (dark colors) groups to the sighted reference group, computer on group's average normalized similarity ratings. Animate agentive and non-agentive (inanimate object) sound verbs are shown separately in different shades of blue. B. Within-group coherence measured as single subjects' correlations to their own group (blind to blind: light; sighted to sighted: dark), using a leave-one-subject-out procedure (See Supplemental Materials for Kendall's W Coherence). Error bars:  $\pm$  standard error of the mean. Significant group differences are marked with p-value. Animate agentive and non-agentive (inanimate object) sound verbs shown separately in different shades of blue.

*gawk, to stare*), brief acts of seeing (e.g. *to peek, to glance, to glimpse*) and acts of generic looking (i.e. *to look, to see, to view*, Fig. 4A).

For light emission verbs, previous linguistic analyses identified intensity and periodicity as two central dimensions of meaning (Faber & Usón, 1999). We therefore asked whether MDS would organize light emission verbs along these dimensions for blind and sighted groups. The verb *to blink* was not included in the original linguistic work, possibly because it has a common meaning unrelated to light emission ("to briefly shut the eyes"; WordNet, 2010) (Faber & Usón, 1999). We therefore excluded *to blink* from this analysis. Consistent with linguistic analyses, the dimensions of intensity and periodicity emerged as the top two for both the sighted and the blind groups. For instance, verbs such as *to flash* and *to blaze* clustered together along the intensity dimension and separated from *to twinkle* and *to glow*. On the periodicity dimension, however, *to flash* and *to twinkle* clustered together and separated from *to blaze* and *to glow* (Fig. 4B). The MDS fit for the light emission verbs was good (low stress values) and comparable across blind and sighted groups (stress value: sighted 0.08, blind 0.07, goodness of fit each group  $p < 0.0001$ ).

### 3.3. The semantic similarity ratings for touch perception and sound emission verbs are more consistent among congenitally blind individuals

As for visual verbs, we used Kendall's W and leave-one-subject out analysis to measure within-group coherence. Blind people's semantic similarity ratings for sound and touch verbs were more consistent across subjects within their own group than those of the sighted

(Fig. 4B, Figure S1, Table S1). This was independently true for animate sound emission (sighted to sighted  $Fisher-z = 0.59$ ,  $SD = 0.16$ ; blind to blind  $Fisher-z = 0.73$ ,  $SD = 0.13$ ; two-sample  $t$ -test across groups  $t(41.53) = 2.33$ ,  $p = 0.02$ ), inanimate sound emission (sighted to sighted  $Fisher-z = 0.39$ ,  $SD = 0.16$ ; blind to blind  $Fisher-z = 0.52$ ,  $SD = 0.11$ ; two sample  $t$ -test across groups  $t(36.19) = 2.59$ ,  $p = 0.01$ ) and touch perception verbs (sighted to sighted  $Fisher-z = 0.58$ ,  $SD = 0.17$ ; blind to blind  $Fisher-z = 0.71$ ,  $SD = 0.1$ ; two-sample  $t$ -test across groups  $t(36.93) = 2.17$ ,  $p = 0.04$ , all within-group effects  $p < 0.0001$ ). Blind participants' ratings for these verb categories were also marginally more correlated with the sighted reference group than were those of the sighted participants (see Supplementary Information for details).

This pattern was specific to sound emission and touch verbs and was not observed for either of the control verb classes (amodal perception verbs: sighted to sighted reference  $Fisher-z = 0.5$ ,  $SD = 0.13$ ; blind to sighted reference  $Fisher-z = 0.45$ ,  $SD = 0.14$ ; two-sample  $t$ -test across groups  $t(43.56) = -0.96$ ,  $p = 0.34$ ; manner of motion verbs: sighted to sighted reference  $Fisher-z = 0.47$ ,  $SD = 0.12$ ; blind to sighted reference  $Fisher-z = 0.47$ ,  $SD = 0.15$ ; two-sample  $t$ -test across groups  $t(43.99) = 0.04$ ,  $p = 0.97$ ).

When comparing touch perception to the amodal control verbs, blind subjects showed higher within-group coherence on the touch verbs relative to the sighted (i.e. were more correlated with their own group (within-group coherence 2 Group(blind, sighted)  $\times$  2 Modality (touch, amodal) repeated measures ANOVAs; main effect of Group  $F(1,44) = 1.94$ ,  $p = 0.17$ , main effect of Modality  $F(1,44) = 30.31$ ,

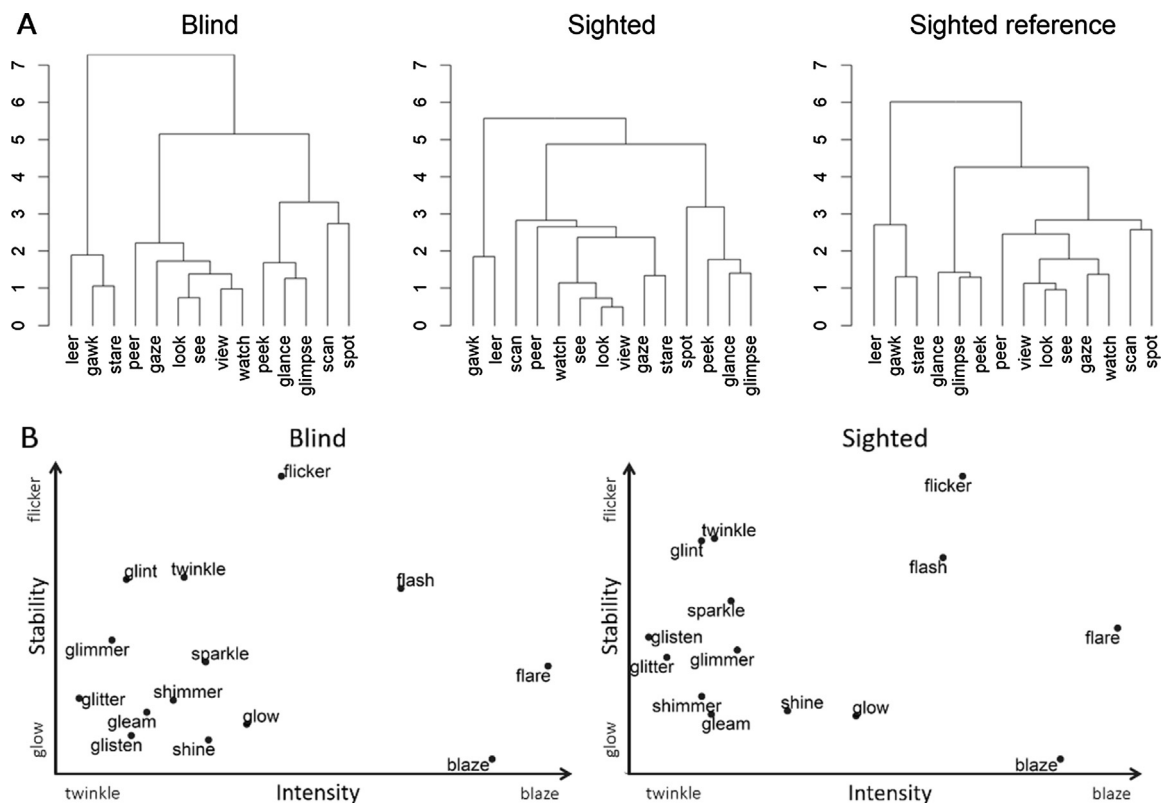


Fig. 4. A. Hierarchical clustering dendrograms for sight perception verbs across groups. B. Light emission verbs MDS analysis, first two dimensions shown.

$p < 0.0001$ , Group  $\times$  Modality interaction  $F(1,44) = 4.93$ ,  $p = 0.03$ ) and were more correlated to the sighted reference group. The group-by-modality interaction did not reach significance for sound emission verbs (within-group coherence 2 Group(blind, sighted)  $\times$  2 Modality (sound, motion) repeated measures ANOVAs; main effect of Group  $F(1,43) = 5.05$ ,  $p = 0.03$ , main effect of Modality  $F(1,43) = 2.49$ ,  $p = 0.12$ , Group  $\times$  Modality interaction  $F(1,43) = 2.79$ ,  $p = 0.1$ ).

Despite these subtle increases in within-group coherence of tactile and sound emission verbs for blind speakers, hierarchical clustering and MDS analyses revealed a qualitatively similar structure and quantitatively similar fits for these verbs across blind and sighted groups (Fig. 5; touch verbs stress value: sighted 0.11, blind 0.09, group-wise goodness-of-fit  $p$ 's  $< 0.0001$ ; sound verbs animate: sighted 0.09, blind 0.08; sound verbs inanimate: sighted 0.13, blind 0.11; group-wise goodness-of-fit  $p$ 's  $< 0.0001$ ). For instance, for both groups touch verbs separated between whole-hands movements (e.g. *to stroke*, *to rub*) and fingertips actions (e.g. *to pinch*, *to prod*). Taken together, these results suggest that blind and sighted individuals rely on shared knowledge when making similarity judgments of touch and sound emission verb, but these meanings are somewhat more likely to be consistent across blind than sighted participants.

## 4. Discussion

### 4.1. Preserved representations of visual verbs in blindness

The present findings reveal similarities between visual verb knowledge among congenitally blind and sighted people. The seminal work of Landau and Gleitman (1985) showed that children who are blind begin to produce and understand the verbs *look* and *see* around the same age as sighted children. Landau and Gleitman (1985) proposed that blind (and sighted) children acquire these meanings partly by relying on language itself i.e. *look* and *see* occur in different syntactic frames (*look at* but not *see at* and *I see that* but not *I look that*) (Gleitman,

1990). The present results extend these previous findings by revealing further richness in the knowledge that blind and sighted people share about visual verbs. First, we find that blind adults treat the modality of perceptual access as a central diagnostic feature of perception verbs; they do not conflate verbs describing visual access to the world (e.g. *to peer*, *to look*) with either tactile (e.g. *to feel*, *to touch*) or amodal verbs (e.g. *to investigate*, *to discover*). A key observation of Landau & Gleitman's original experiment was that the blind 4-year-old, Kelly, raised her hands expecting to examine something when instructed to *look*. Kelly interpreted the instruction to *look* as 'observe with the hands', when *look* referred to herself. Unlike the experiment with Kelly, the current study presented verbs without specifying whether the agent was sighted or blind. In this unspecified context, blind and sighted adults alike distinguish verbs such as *look*, *stare* and *see* from amodal and tactile verbs. Using feature generation, Lenci et al. (2013) similarly found that blind people generate visual features when asked the meaning of visual verbs. Thus, although speakers are able to flexibly and appropriately apply verbs such as *look* and *see* to blind agents, the present data suggest that when no agent is specified, visual verbs imply a sighted agent even for blind speakers.

We further find that the within-category similarity structure of visual verbs is preserved in blindness. Blind people's ratings of visual verbs were not more noisy, or heterogeneous than those of the sighted, unlike the case of color similarity judgments, which are more variable across blind individuals (Saysani et al., 2018; Marmor, 1978; Saysani et al., 1992). Blind individuals represent the temporal structure of visual perception verbs (e.g. distinguishing *glance* vs. *stare*) and distinguish light events along dimensions of temporal frequency (*glow* vs. *twinkle*) and intensity (*glow* vs. *blaze*). In sum, blind and sighted people share detailed knowledge pertaining to acts of visual perception and events of light emission.

One question, however, is whether the reported pair-wise similarity judgments provide an accurate measure of word meanings or concepts. Perhaps, blind participants performed the task based on just the



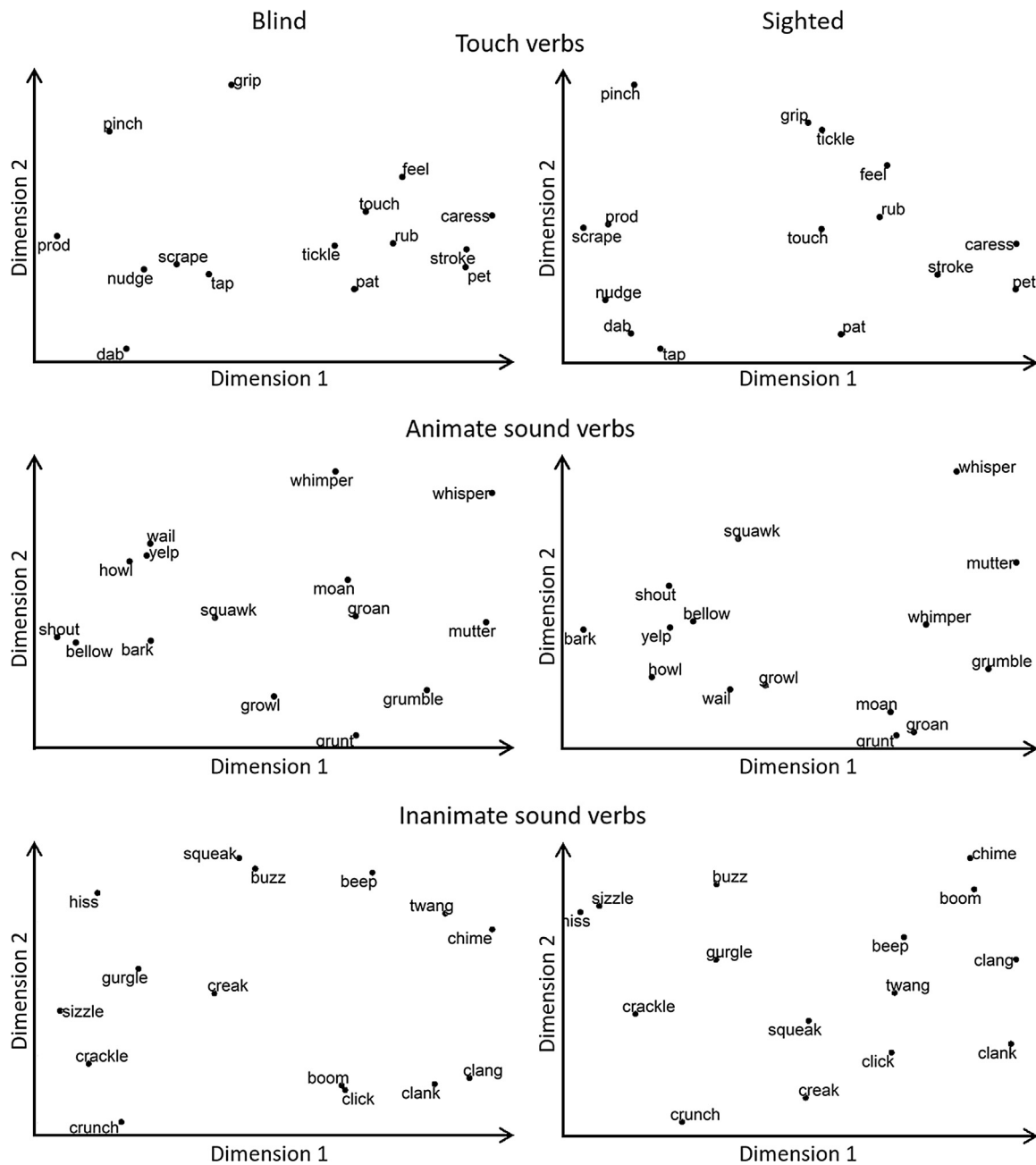


Fig. 5. MDS results for touch, animate/agentive sound and inanimate-object sound emission verbs MDS for blind and sighted groups. Top two dimensions shown. Note that dimension 2 was rotated for touch verbs, sighted group and inanimate sound verbs, sighted group.

distribution of co-occurrence of words in others’ speech, not a representation of the words’ meaning. We think this explanation is unlikely for a number of reasons. First, semantic similarity judgments of the kind used in the current study are strongly correlated with performance on other conceptual tasks, such as categorization (is a robin a bird?) (e.g. Rips et al., 1973). Second, as noted above, smaller scale studies using language production, and explicit definitions, have also found that blind individuals use visual words in meaningful ways during communication (Landau & Gleitman, 1985; Lenci et al., 2013). Third, blind and sighted people recruit the same brain regions when making inferences based on sentences with visual perception verbs (Bedny, Pascual-Leone, & Saxe, 2009; Koster-Hale, Bedny, & Saxe, 2014).

The present data also provides evidence that the semantic similarity task used here is sensitive to subtle between-group differences in meaning knowledge as evidenced by blind individuals’ somewhat more

consistent judgments of tactile perception and sound emission verbs. As a group, blind adults’ ratings for these verbs are indistinguishable from those of the sighted (i.e. blind and sighted speakers use similar criteria to distinguish *stroke* from *prod*) but blind people’s responses are less heterogeneous across participants. Why might this be the case? One possible interpretation is that superior abilities in touch and hearing among blind individuals translate to more precise meanings for touch and sound verbs. Indeed, blind individuals outperform the sighted on some tactile and auditory tasks (Gougoux et al., 2004; Van Boven, Hamilton, Kauffman, Keenan, & Pascual-Leone, 2000). Such an interpretation of the current data seems unlikely to us, however, for multiple reasons. First, blind individuals only outperform sighted individuals on a subset of tactile and auditory perceptual tasks and the benefits are subtle. For example, blind individuals are better at perceiving Braille-like patterns but not other types of tactile stimuli (Grant, Thiagarajah, & Sathian, 2000). There are even some auditory tasks, such as localization

in the vertical plane, on which sighted individuals perform better (Zwiers, Van Opstal, & Cruysberg, 2001). It is unlikely that such subtle and uneven sensory enhancements in blindness cause lexical differences. Furthermore, if total absence of first-person visual experience does not make visual verb meanings more noisy in blind adults, it is unlikely that subtle improvements in tactile or auditory experience would make them less noisy.

We hypothesize instead that differences in communicative relevance account for differences in semantic similarity judgment consistency across blind and sighted speakers. Since blind individuals live among a sighted majority, being blind oneself does not substantially reduce the frequency of encountering visual verbs. By contrast, being blind might make one more likely to use verbs of sound emission and touch perception during communication. Blind people, for example, may be more likely than the sighted to ask to touch or hold something, or to describe an event's auditory properties. This account is speculative and requires future testing. Measuring the frequency of sound and touch verbs in naturalistic speech and writing of individuals who are blind could provide relevant evidence. Regardless of the underlying cause, the present findings suggest that semantic similarity judgments can in principle measure between-group differences in meaning that result from blindness, making the observation that the semantic similarity judgments for visual verbs are not altered by congenital blindness all the more compelling.

At the same time, the reported results are only a partial measure of what blind and sighted speakers know about events of light and visual perception and therefore are not the final word on whether there are knowledge differences across these populations. There are a number of reasons why semantic similarity judgments of the kind used in the current study do not provide direct access to the multidimensional “psychological space” of conceptual representations (Medin, Goldstone, & Gentner, 1993; Shepard, 1987; Tenenbaum & Griffiths, 2001). For one, semantic similarity judgments are highly sensitive to context (e.g. Gauker, 1994; Goodman, 1972; Goldstone, 1994). People's similarity judgments shift rapidly and flexibly. Two objects that seem initially wildly dissimilar (e.g. children and jewelry) can easily be judged as similar when given the right frame (“objects to rescue from a burning house”, Barsalou, 1983). Gray is more similar to white than black, when the context is hair color, but more similar to black than white, when the context is clouds (Goldstone, 1994; Medin & Shoben, 1988). Indeed, the context may be just the order of presentation of the pair of words: “To say that surgeons are like butchers means something different than to say butchers are like surgeons” (Medin et al., 1993). Similarly, Ross and Murphy (1999) found that college students recognize two orthogonal ways to organize food categories: by taxonomy (e.g. milk and ice cream are both dairy foods) and by social context (e.g. milk and bagels are both breakfast foods). When simply asked about the “similarity” of two foods, participants tended to prioritize taxonomical categories, judging milk more similar to ice cream; but when asked to make inferences about social behaviours (e.g. inclusion in a novel ritual), participants made predictions based on social scripts. In sum, there is clear evidence that people know more about concepts and words than what is captured by a particular semantic similarity judgment task.

We therefore do not interpret the present similarity judgments as directly revealing the representational space of the meaning of visual verbs, in either sighted or blind individuals. Instead, similarity judgments provide a sensitive but incomplete estimate of people's knowledge of a domain. Thus, the best explanation of our results is that sighted and blind individuals share both (i) relevant knowledge of the meanings of “visual verbs”, and (ii) common pragmatics, that lead them to interpret the request for similarity judgments in terms of the relevant respects for this domain (e.g. modality, temporal duration, etc).

#### 4.2. Some open questions

The present results demonstrate that blind individuals know

important aspects of visual verb meanings, but do not speak to how these are acquired. How do blind individuals learn the meanings of visual verbs? Language itself is likely a rich source of information. The meanings of visual verbs may be partly inferred from the meanings of the phrases in which they occur. Landau & Gleitman argued that blind children use sentence frames to distinguish between *look* and *see* (1985). Analogously, hearing “the light flashed on and off” as opposed to the “light glowed” might provide clues to the temporal structure of light events. Words that occur in similar linguistic environments have more similar meanings (Landauer & Dumais, 1997). Blind learners could use the meanings of words they already know, their interpretation of the discourse as well as social and pragmatic cues to constrain hypotheses about visual words (Bloom, 2002; Clark, 1987; Frank & Goodman, 2014; Markman & Wachtel, 1988; Ouyang, Boroditsky, & Frank, 2016; Tomasello & Barton, 1994). For example, when a blind individual hears a sighted speaker comment on *seeing* a *glowing* star in the night sky, she might infer that for sighted people *seeing* occurs at a distance and that *glowing* can be ascertained through vision but not through audition, since she herself cannot observe the glowing of the star. When she hears someone complain of being *stared at* throughout lunch, she might infer that staring is something that can last all lunch long. In this way linguistic, social and pragmatic information, together with a shared innate endowment for processing it, act to align the minds of sighted and blind speakers. Understanding exactly how this occurs is an important goal for future research that could be attained, in part, by studying language acquisition in blind children.

A further open question concerns how knowledge of visual verbs among blind individuals compares to knowledge about other visual domains. Prior studies suggest that visual verbs are not the only part of visual knowledge that is preserved in blindness. As noted in the introduction, even blind children know that colors are physical properties that are perceptible only with the eyes (Landau & Gleitman, 1985). Blind adults' similarity judgments on colors produced a color wheel qualitatively similar to that of sighted adults (e.g. blue is more similar to green than red). However, relative to the sighted, there is higher variability in color similarity knowledge across blind individuals (Marmor, 1978; Sayani et al., 2018; Shepard & Cooper, 1992). Moreover, blind individuals are less likely than the sighted to use color during semantic similarity judgments (Connolly, Gleitman & Thompson-Schill, 2007). Why are the meanings of visual verbs less variable than the similarities among colors? One possibility is that some information, such as between color similarity, is less inferentially relevant and therefore blind individuals are less likely to learn it. Connolly et al. (2007) suggest that for blind individuals “strawberries are known to be red, [but] nothing follows in terms of the usefulness of this fact in reasoning about strawberries.” Whether an agent is *staring* or *peeking* might license more inferences (e.g. about what the agent knows), than knowing the color of an object. Another possibility is that some visual information is more easily accessible through language. In future work, computational models could be used to ask which vision-related information is most available in text. We would then be in a position to compare and contrast what is available and what blind speakers actually learn.

Further work is needed to fully characterize blind and sighted people's knowledge about vision and light. The stimuli used here were single words and verbs in particular. The meanings of single words are necessarily general and flexible, therefore, perhaps most likely to be robust to changes in our idiosyncratic life histories. When words are combined into phrases, the generated meanings are more than the sum of their parts, and the inferences that follow additionally depend on real-world knowledge. For example, the inferences one makes based on the phrase “Abigail glanced at Leo's face across the room” depend on an understanding of vision that goes beyond the meaning of *glance*. Does Abigail know the color of Leo's eyes? Does she have information about his mood? Whether he's hungry? Does she know whether Leo is wearing a hat? The color of his shoes? Future work comparing blind and sighted

people's inferences about the visual experiences of others would reveal further insight into the contribution of vision to knowledge acquisition.

## 5. Conclusions

The present findings reveal a rich set of knowledge about vision and light that is shared among sighted and blind individuals. These results provide a compelling illustration of the shared nature of meaning and its resilience to dramatic change in first-person sensory histories.

## Acknowledgements

We would like to thank members of the blind community for generously giving of their time to make this project possible. We thank William Johnson for his assistance in data collection and gratefully acknowledge the David and Lucile Packard Foundation for funding this project.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2019.03.017>.

## References

- Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, 11(3), 211–227.
- Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. *Concepts and conceptual development: Ecological and intellectual factors in categorization*, 10139.
- Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. *Theories of Memory*, 1, 29–31.
- Barsalou, L. W., & Sewell, D. R. (1984). Constructing representations of categories from different points of view.
- Bedny, M., Pascual-Leone, A., & Saxe, R. R. (2009). Growing up blind does not change the neural bases of Theory of Mind. *Proceedings of the National Academy of Sciences*, 106(27), 11312–11317.
- Berkeley, G. (1948). An essay toward a new theory of vision, 1709. In W. Dennis (Ed.), *Century psychology series. Readings in the history of psychology* (pp. 69–80). East Norwalk, CT, US: Appleton-Century-Crofts.
- Berkeley, G. (1732). An Essay towards a New Theory of Vision. (Based on the fourth edition, London, 1732). Edited by David R. Wilkins, Dublin, December 2002.
- Bimler, D., & Kirkland, J. (2007). Constructing personality maps, mapping personality constructs: Multidimensional scaling recovers the big five factors from internal and external structure. *The Spanish Journal of Psychology*, 10(1), 68–81.
- Bloom, P. (2002). Mindreading, communication and the learning of names for things. *Mind and Language*, 17, 37–54.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990.
- Chan, A. S., Butters, N., Paulsen, J. S., Salmon, D. P., Swenson, M. R., & Maloney, L. T. (1993). An assessment of the semantic network in patients with Alzheimer's disease. *Journal of Cognitive Neuroscience*, 5(2), 254–261.
- Clark, J. M. (1987). Understanding pictures and words: Comment on Potter, Kroll, Yachzel, Carpenter, and Sherman (1986). *Journal of Experimental Psychology: General*, 116, 307–309.
- Cannolly, A. C., Gleitman, L. R., & Thompson-Schill, S. L. (2007). Effect of congenital blindness on the semantic representation of some everyday concepts. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 8241–8246.
- Cutsforth, T. D. (1932). The unreality of words to the blind. *Teachers Forum*, 4, 86–89.
- Cutsforth, T. D. (1951). *The blind in school and society*. New York: American Foundation for the Blind.
- De Leeuw, J., & Mair, P. (2011). Multidimensional scaling using majorization: SMACOF in R.
- De Leeuw, J., & Heiser, W. J. (1977). Convergence of correction matrix algorithms for multidimensional scaling. *Geometric Representations of Relational Data*, 735–752.
- Faber, P. B., & Usón, R. M. (1999). Constructing a lexicon of English verbs (Vol. 23). Walter de Gruyter.
- Frank, M. C., & Goodman, N. D. (2014). Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology*, 75, 80–96.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition: A Journal of Developmental Linguistics*, 1(1), 3–55.
- Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, 52(2), 125–157.
- Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects*. New York: Bobbs-Merrill.
- Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., & Belin, P. (2004). Neuropsychology: Pitch discrimination in the early blind. *Nature*, 430(6997), 309.
- Grant, A. C., Thiagarajah, M. C., & Sathian, K. (2000). Tactile perception in blind Braille readers: A psychophysical study of acuity and hyperacuity using gratings and dot patterns. *Perception & Psychophysics*, 62, 01–312.
- Harrell Jr., F. E., & Dupont, C. (2014). Hmisc: Harrell Miscellaneous, 2011. URL <http://CRAN.R-project.org/package=Hmisc>. R package version, 3-9.
- Hobbes, T. (1641/1984). Meditations on first philosophy, objections with replies. In Robert Stoothoff, & Douglas Murdoch (Eds.). *The Philosophical Writings of Descartes*. Cambridge, UK: Cambridge University Press.
- Hume, D. (1739/1978). *A treatise of human nature*. Oxford, UK: Oxford University Press.
- Hutchinson, J., & Lockhead, G. R. (1977). Similarity as distance: A structural principle for semantic memory. *Journal of Experimental Psychology: Human Learning and Memory*, 3(6), 660.
- Koster-Hale, J., Bedny, M., & Saxe, R. (2014). Thinking about seeing: Perceptual sources of knowledge are encoded in the theory of mind brain regions of sighted and blind adults. *Cognition*, 133(1), 65–78.
- Landau, B., & Gleitman, L. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211.
- Lenci, Baroni, M., Cazzolli, G., & Marotta, G. (2013). BLIND: A set of semantic feature norms from the congenitally blind. *Behavioral Research Methods*, 45, 1218–1233.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Locke, J., & Nidditch, P. H. (2011). *An essay concerning human understanding*. Oxford: Clarendon Press.
- Lopez, A., Atran, S., Coley, J. D., Medin, D. L., & Smith, E. E. (1997). The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology*, 32(3), 251–295.
- Mair, P., de Leeuw, J., & Groenen, P. J. (2015). Multidimensional scaling in R: smacof. URL: <https://cran.r-project.org/web/packages/smacof/vignettes/smacof.pdf>.
- Majid, A., Roberts, S. G., Cilissen, L., Emmorey, K., Nicodemus, B., O'Grady, L., ... Shayan, S. (2018). Differential coding of perception in the world's languages. *Proceedings of the National Academy of Sciences*, 115(45), 11369–11376.
- Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20, 121–157.
- Marmor, G. S. (1978). Age at onset of blindness and the development of the semantics of color names. *Journal of Experimental Child Psychology*, 25(2), 267–278.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, 100(2), 254.
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, 20(2), 158–190.
- Murphy, G. (2004). *The big book of concepts*. MIT press.
- Murtagh, Fionn, & Legendre (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *Journal of Classification*, 31, 274–295.
- Ouyang, L., Boroditsky, L., & Frank, M. C. (2017). Semantic coherence facilitates distributional learning. *Cognitive Science*, 41, 855–884.
- Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 12(1), 1–20.
- Roberts, J. S., & Wedell, D. H. (1994). Context effects on similarity judgments of multidimensional stimuli: Inferring the structure of the emotion space. *Journal of Experimental Social Psychology*, 30(1), 1–38.
- Rosel, J., Caballer, A., Jara, P., & Oliver, J. C. (2005). Verbalism in the narrative language of children who are blind and sighted. *Journal of Visual Impairment & Blindness*, 99, 413–425.
- Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology*, 38, 495–553.
- Rumelhart, D. E., & Abrahamson, A. A. (1973). A model for analogical reasoning. *Cognitive Psychology*, 5(1), 1–28.
- Sayani, A., Corballis, M. C., & Corballis, P. M. (2018). Color envisioned: Concepts of color in the blind and sighted. *Visual Cognition*, 26(5), 382–392.
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237(4820), 1317–1323.
- Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive Psychology*, 1(1), 1–17.
- Shepard, R. N., & Cooper, L. A. (1992). Representation of colors in the blind, color-blind, and normally sighted. *Psychological Science*, 3(2), 97–104.
- Shepard, R. N., Kilpatrick, D. W., & Cunningham, J. P. (1975). The internal representation of numbers. *Cognitive Psychology*, 7(1), 82–138.
- Suzuki, R., & Shimodaira, H. (2006). Pvcust: An R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, 22(12), 1540–1542.
- Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24(4), 629–640.
- Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, 30(5), 639.
- Van Boven, R. W., Hamilton, R. H., Kauffman, T., Keenan, J. P., & Pascual-Leone, A. (2000). Tactile spatial resolution in blind Braille readers. *Neurology*, 54(12), 2230–2236.
- Winter, B., Perlman, M., & Majid, A. (2018). Vision dominates in perceptual language: English sensory vocabulary is optimized for usage. *Cognition*, 179, 213–220.
- Zwiers, M. P., Van Opstal, A. J., & Cruysberg, J. R. M. (2001). A spatial hearing deficit in early-blind humans. *Journal of Neuroscience*, 21(9), RC142 RC142-RC142.