

The final publication is available at Springer via <http://dx.doi.org/10.3758%2Fs13428-013-0323-4>

BLIND: A set of semantic feature norms from the congenitally blind
Alessandro Lenci¹, Marco Baroni², Giulia Cazzolli² and Giovanna Marotta¹

¹Department of Philology, Literature, and Linguistics, University of Pisa

²Center for Mind/Brain Sciences, University of Trento

Authors Note

This research was carried out within the project “Semantic representations in the language of the blind: linguistic and neurocognitive studies” (PRIN 2008CM9MY3), funded by the Italian Ministry of Education, University and Research. We thank Alessandro Fois, Margherita Donati and Linda Meini, for help with data collection, Sabrina Gastaldi who collaborated to the norms collection and annotation, Giovanni Cassani for the English translation of concept properties, and Daniela Bonino for the medical interviews. We thank the Unione Italiana Ciechi (National Association for the Blind) for their collaboration.

Correspondence concerning this article should be addressed to Alessandro Lenci, Dipartimento di Filologia, Letteratura e Linguistica, via Santa Maria 36, 56126 PISA, Italy, alessandro.lenci@ling.unipi.it

Abstract

Feature-based descriptions of concepts produced by subjects in a property generation task are widely used in cognitive science to develop empirically grounded concept representations and to study systematic trends in such representations. This article introduces BLIND, a collection of parallel semantic norms collected from a group of congenitally blind Italian subjects and comparable sighted subjects. The BLIND norms comprise descriptions of 50 nouns and 20 verbs. All the materials have been semantically annotated and translated into English, to make them easily accessible to the scientific community. The article also presents a preliminary analysis of the BLIND data that highlights both the large degree of overlap between the groups and interesting differences. The complete BLIND norms are freely available and can be downloaded from http://sesia.humnet.unipi.it/blind_data.

Keywords: blind, feature norms, semantic memory, feature production experiment, feature type annotation

Introduction

The effect of blindness on the organization and structure of conceptual representations has always been regarded as crucial evidence to understand the relationship between sensory-motor systems and semantic memory, as well as the link between language and perception. Researchers have been struck by the close similarity between the language of the congenitally blind and sighted, even for those areas of the lexicon that are directly related to visual experience, such as color terms or verbs of vision. In a multidimensional scaling analysis performed by Marmor (1978) with similarity judgments about color terms, the similarity space of the congenitally blind subjects closely approximates Newton's color wheel and judgments by sighted control participants. Therefore, she concludes that knowledge of color relations can be acquired without first-hand sensory experience. Zimler and Keenan (1983) also found no significant differences between blind and sighted in a free-recall task for words grouped according to a purely visual attribute, such as the color red (e.g., *cherry* and *blood*). Kelli, the congenitally blind child studied by Landau and Gleitman (1985), was able to acquire impressive knowledge about color terms, including the constraints governing their correct application to concrete nouns, without overextending them to abstract or event nouns. Kelli also properly used other vision-related words, like the verbs *look* and *see*, though her meaning of *look* seemed to apply to haptic explorations. Apart from some delay in the onset of speech, Kelli showed normal language development, with her lexicon and grammar being virtually undistinguishable from the ones of sighted children by the age of three. The common interpretation of these data is that congenitally blind people possess substantial knowledge about the visual world derived through haptic, auditory and linguistic input.

These linguistic similarities notwithstanding, the blind still lack the *qualia* associated with the perception of visual features, such as colors. Whether this experiential gap leads to different conceptual representations in the sighted and the

blind is an important aspect of the current debate on semantic memory and its relationship with sensory systems. According to the embodied cognition approach, concepts are made of inherently modal features and concept retrieval consists in reactivating sensory-motor experiences (Barsalou, 2008; De Vega, Glenberg, & Graesser, 2008; Gallese & Lakoff, 2005; Pulvermüller, 1999). If the visual experience of a red cherry is part and parcel of the concept of cherry in the sighted, the embodied view therefore predicts that the concept of cherry in the congenitally blind should be substantially different, given the lack of the visual component. Most neuroscientific evidence instead reveals a strong similarity between blind and sighted individuals, thereby supporting the view that conceptual representations are more abstract than the ones assumed by embodied models (Cattaneo & Vecchi, 2011). In a fMRI study with congenitally blind subjects, Pietrini et al. (2004) found responses to category-related patterns in the visual ventral pathway, like in the sighted, and thus argued that this cortical region actually contains more abstract, “supramodal” representations of object form (Pietrini, Ptito, & Kupers, 2009). According to another view, closely related to the supramodality hypothesis, semantic memory is formed by modality-independent representations (Bedny, Caramazza, Pascual-Leone, & Saxe, 2012; Mahon & Caramazza, 2008). Concepts are stored in non-perceptual brain regions and consist of abstract, “symbolic” features. This model also predicts that the concepts of congenitally blind subjects are highly similar to those of the sighted. The same abstract representation of a cherry could, in fact, be created using information coming from language and from senses other than vision. Mahon, Anzellotti, Schwarzbach, Zampini, and Caramarazza (2009) indeed show that sighted and blind individuals do not differ at the level of neural representation of semantic categories in the ventral visual pathway. Bedny et al. (2012) find parallel results with fMRI activations for action verbs in the left middle temporal gyrus, consistently with other evidence supporting the close similarity of the neural representations of actions in blind and sighted subjects (Noppeney, Friston & Price, 2003; Ricciardi et al., 2009).

Linguistic and neuroscientific evidence converge in revealing close similarities between conceptual representations in the blind and the sighted, which depend on the possibility of the former to build rich and detailed representations of the world by combining linguistic and non-visual sensory information. Still, the picture is far from being unequivocally clear. Even the similarities between color spaces in congenitally blind subjects are not without controversy. For instance, Shepard and Cooper (1992) found important differences between the color spaces of sighted and congenitally blind subjects, differently from Marmor (1978). Connolly, Gleitman, and Thompson-Schill (2007) also showed that the lack of visual experience of colors has significant effects for conceptual organization in blind subjects. They collected implicit similarity judgments in an odd-man-out task about two categories of concepts, “fruits and vegetables” and “household items”. Cluster analysis of the similarity judgments revealed a major overlap between the blind and sighted similarity spaces, but significant differences for clusters of the “fruit and vegetables” category for which color is a diagnostic property (i.e., critical to identify the exemplars of that category, such as being yellow for a banana). Even for blind subjects with good knowledge of the stimulus color, such information does not appear to affect the structure of the similarity space. The hypothesis by Connolly et al. is that such contrast stems from the different origin of color knowledge in the two groups. In the congenitally blind, color knowledge is “merely stipulated”, because it comes from observing the way color terms are used in everyday speech, while in the sighted it is an immediate form of knowledge derived from direct sensory experience, and used to categorize new exemplars. This results in a different organization of concepts in which color enters as a critical feature. Similar conclusions are also drawn by Marques (2009), who tested early onset blind people and sighted subjects on a recall task of word triads sharing visual properties. These data do not necessarily undermine “abstract” views of concepts (cf. Bedny et al., 2012), but they show that semantic

features not properly grounded on sensory systems might have a different status in concepts.

With the debate on the effects of visual deprivation on conceptual representations still open, and playing a fundamental role in answering some central questions in cognitive neuroscience, the field can benefit from public data sets enabling empirical studies of blind conceptualization. To our knowledge, however, such resources are missing. To partially address this gap, we introduce here BLIND (*BLind Italian Norming Data*), a set of semantic feature norms collected from Italian congenitally blind and sighted individuals. Semantic features (also known as “properties”) play a key role in most models of concepts (Murphy, 2002; Vigliocco & Vinson, 2007). Feature norms are collected in a property generation task: Subjects are presented with concept names on a blank form, and are asked to write down lists of properties to describe the entities the words refer to. Collected properties then undergo various kinds of processing to normalize the produced features, and typically to classify them with respect to a given set of semantic types. Responses are pooled across subjects to derive average representations of the concepts, using subjects’ production frequency as an estimate of feature salience. In addition, other statistics useful to characterize the feature distribution among concepts as well as various measures that enrich the norms, such as concept familiarity ratings, word frequency counts, etc., are provided.

After Garrard, Lambon Ralph, Hodges, and Patterson (2001) made available their norming study on 64 living and non-living concepts, other feature norms for English have been freely released. The norms by McRae, Cree, Seidenberg, and McNorgan (2005) consist of features for 541 living and non-living basic-level noun concepts, while Vinson and Vigliocco (2008) collected norms for object and event nouns, as well as verbs for a total of 456 words. De Deyne et al. (2008) extended the work by Ruts et al. (2004) and published feature norms for 425 Dutch nouns belonging to 15 categories (these also include categories such as professions and

sports, that do not typically appear in other norming studies). As for Italian, Kremer and Baroni (2011) asked German and Italian native speakers to norm 50 noun concepts belonging to 10 categories, while Montefinese, Ambrosini, Fairfield, and Mammarella (2012) collected feature norms for 120 artifactual and natural nouns. Typically, feature norms are elicited by presenting stimuli as single words out of context (a short context might be provided to disambiguate homograph words). Frassinelli and Lenci (2012) instead carried out a small-scale norming study in which English nouns appeared in different types of visual and linguistic contexts, with the goal of investigating the effect of context variation on feature production.

Feature norms are neutral with respect to the format of underlying conceptual representations (e.g., semantic networks, frames, perceptual simulators, prototypes, etc.). Produced features are not to be taken as literal components of concepts, but rather as an overt manifestation of representations activated by subjects during the task (MacRae et al., 2005). When presented with a stimulus, the corresponding concept is accessed and “scanned” by subjects to produce relevant properties. For instance, Barsalou (2003) assumes that, when they generate feature lists, subjects use modal perceptual simulation of the word referent (Barsalou, 1999), while Santos, Chaigneau, Simmons, and Barsalou (2011) claim that feature norms reflect both sensory-motor information re-enacted as conceptual simulation, and linguistic associations.

Feature norms have some notorious limits widely discussed in the literature (McRae et al., 2005), the major of which is the verbal nature of the features. The problem is that not all aspects of meaning are equally easy to describe with language. For instance, a subject could be acquainted with the particular gait of an animal, without being able to find a verb to express it, or be aware of some parts of an object, without knowing their name. Therefore, there is always the risk that important concept features are underrepresented in the norms. Moreover, subjects tend to produce features that distinguish a concept from others in the same category, thereby

resulting in very general features to be also underproduced. For instance, when describing a canary subjects focus on properties that distinguish it from other birds (e.g., color, size, etc.), rather than on features that are common to all birds or animals (e.g., it has two eyes, it is alive, etc.). De Deyne et al. (2008) tried to address this problem by collecting also applicability judgments of features to superordinate categories of the normed concepts. Finally, subjects often produce very complex linguistic descriptions that fuse multiple features or correspond to complex arrays of knowledge (e.g., *birds migrate towards warm areas in winter*).

Their limits notwithstanding, feature norms provide significant approximations to concept structure and are used to develop empirically grounded models of semantic memory, complementing the evidence coming from behavioral and neuro-exploratory experiments. In fact, feature norms have been applied to develop accounts of a wide range of empirical phenomena, from categorization to semantic priming, from conceptual combination to category-specific deficits (Cree & McRae, 2003; Garrard et al., 2001; Vigliocco, Vinson, Lewis, & Garrett, 2004; Vinson, Vigliocco, Cappa, & Siri, 2003; Wu & Barsalou, 2009). On the computational modeling side, feature norms are used as behavioral benchmarks for distributional semantic models (Baroni & Lenci, 2008; Baroni, Barbu, Murphy & Poesio, 2010; Baroni & Lenci, 2010; Riordan & Jones, 2010), to develop “hybrid” representations integrating linguistic and experiential features (Andrews, Vigliocco, & Vinson, 2009; Steyvers, 2010), and in neurocomputational models of semantic decoding of fMRI activations (Chang, Mitchell, & Just, 2010).

The main purpose of this paper is to introduce the BLIND norms as a new, freely available data resource to explore the distribution of conceptual features in congenitally blind and sighted persons. The complete BLIND norms may be explored and downloaded from http://sesia.humnet.unipi.it/blind_data. We hope that the BLIND data will allow researchers to gain new evidence on the role of visual experience in the organization of concepts and to understand the salience and

distribution of semantic feature types. We discuss next the methodology used to collect BLIND and we carry out a comparison of the sighted subject data in it with the Kremer and Baroni (2011) norms, suggesting that our data are fully comparable with those in the earlier conceptual norm literature, despite some differences in the nature of the subject pool and data collection methodology. We then describe how the norms can be accessed, and we present a few preliminary analyses of feature type distributions in BLIND as a concrete example of how this resource can be used. The emerging picture reveals similarities between the blind and the sighted, laying side by side with very significant differences.

The BLIND norms

The BLIND norms were collected with a feature generation task inspired by previous norming studies, in particular McRae et al. (2005), Vinson and Vigliocco (2008) and Kremer and Baroni (2011). However, there are some major elements of novelty that differentiate BLIND from existing feature norms: (a) For the first time semantic features were produced by congenitally blind persons, together with a comparable sample of sighted individuals; (b) Subjects described the words orally, thereby gaining spontaneity and overcoming the stricture of written form filling; (c) Stimuli include both concrete and non-concrete nouns and verbs.

Participants

The norming study was conducted on 48 Italian subjects, 22 congenitally blind (10 females, 2 left-handed) and 26 sighted (13 females, 1 left-handed). Blind volunteers were recruited with the help of local sections of the Unione Italiana Ciechi (National Association for the Blind), which promotes activities and organizes education and professional courses within the blind community. The districts of provenience of the blind subjects were Tuscany (12), Liguria (5), and Sardinia (5). Their age was in the range of 20 to 73 years and the average age was 47.2 (SD 16.5). Education levels ranged from junior high school to a master degree. Most subjects finished their studies after graduating from high school (11). The majority of blind

subjects were switchboard operators (14). The others were teachers, students or they had already retired. Some had close-to-normal mobility, while others had a very low degree of autonomy. All blind subjects were proficient Braille readers. The 26 control sighted subjects were selected to match blind subjects as close as possible with respect to age, gender, residence, education, and profession. The districts of provenience of the sighted subjects were Tuscany (16), Liguria (5), and Sardinia (5). Their age was in the range of 18 to 72 years and the average age was 45.1 (SD 16.8). It is worth noting that our participants are much older than subjects used in similar norming studies, typically university students. This is due to the fact that in Italy, like in other developed countries, congenital blindness is steadily reducing, and most of the people affected by this impairment are in their 40s or older. All subjects gave informed consent and received a small expense reimbursement.

Participants were native speakers of Italian and none of them suffered from neurological or language disorders. Blind subjects received a medical interview to check their status of total, congenital blindness. Main causes of blindness were congenital glaucoma, retinopathy of prematurity, and congenital optic nerve atrophy retinopathy. Before the experiment, all subjects performed a direct and backward digit span test, revealing normal working memory (sighted average direct: 6.8, SD: 1; blind average direct: 7.6; SD: 1.2; sighted average backward: 5.3, SD: 1; blind average backward: 5.2; SD: 1.5). Notice that the blind direct digit span was significantly higher than the sighted one, according to a Wilcoxon test ($W = 408.5, p < 0.01$).

Stimuli

The normed concepts correspond to 50 Italian nouns and 20 Italian verbs equally divided among 14 classes (see Appendix A for the complete list of Italian stimuli, their English translation, and classes). The nouns consist of 40 concrete and 10 non-concrete concepts. Concrete nouns cover various living and non-living classes, most of which were already targeted by previous norming studies (Kremer & Baroni, 2011; McRae et al., 2005) or by experiments with blind subjects (Connolly et

al., 2007). Special care was taken to include nouns for which visual features are particularly salient and distinctive (e.g., stripes for *zebra*, color for *banana*, long neck for *giraffa* “giraffe”, etc.). Non-concrete nouns are divided into emotions (e.g., *gelosia* “jealousy”) and abstract nouns expressing ideals (e.g., *libertà* “freedom”). Fifteen verbs express modes of visual (e.g., *scorgere* “to catch sight of”), auditory (e.g., *ascoltare* “to listen”), and tactile (e.g., *accarezzare* “to stroke”) perception; five verbs express abstracts events or states (e.g., *credere* “to believe”).

Stimulus frequencies as reported in the norms were estimated on three Italian corpora: itWac (Baroni, Bernardini, Ferraresi & Zanchetta, 2009), Repubblica (Baroni et al., 2004), and a dump of the Italian Wikipedia (ca. 152 million word tokens). As it can be expected from the composition of the above corpora (newspaper articles for Repubblica and web pages for itWac and Wikipedia), birds, fruits, mammals, vegetables and tools have much lower frequencies than vehicles, locations and non-concrete nouns. The mean frequency of auditory and abstract verbs is also higher than for visual and tactile verbs.

Procedure

The subjects listened to the randomized noun and verb stimuli presented on a laptop with the PsychoPy software. In order to reduce fatigue effects, the experiment was split in two sessions, with the second session occurring no less than 10 days after the first one. At the beginning of each session, subjects listened to the task instructions (their English translation is reported in Appendix B). Subjects were asked to orally describe the meaning of the Italian words. In order to reduce the possible impact of free associations, subjects were instructed not to hurry, to think carefully about the meaning and the properties of the stimuli, and to express them with short sentences and phrases. An example of how a noun and verb could be described were also provided. After listening to the instructions, the subjects performed a short trial session, norming two nouns and one verb. Neither the instruction nor the trial

concepts belonged to the classes of the test stimuli. Both the instructions and the trial could be repeated, until the task was perfectly clear.

While we instructed the subjects not to hurry, to avoid too much lingering on specific concepts, we fixed a maximum time limit of one minute per stimulus, after which subjects heard a “beep” and the next stimulus was presented. We established the time limit based on our previous experience in collecting conceptual norms, having observed that one minute is typically plenty of time for a full concept description even in the slower written production modality. Subjects could decide to pass to the next stimulus by pressing the spacebar at any time. If a word was not understood clearly, subjects could listen to it again, by pressing a computer key. Halfway through the task, subjects took a five minute break. The subjects listened to the stimuli over headphones and produced their descriptions orally using a microphone. The descriptions were recorded as WAV files.

After completing the second session, subjects performed a familiarity rating test on the experimental stimuli. They were asked to rate their level of experience with the object (event or action) to which the noun (verb) refers to on a scale of 1 to 3, with 1 corresponding to little or no experience, and 3 to great experience. The average per-concept mean familiarity is 2.34 (SD: 0.54) for blind and 2.32 (SD: 0.45) for sighted subjects (difference not significant according to a paired Wilcoxon test). The experiments and tests were conducted at the Laboratory of Phonetics of the University of Pisa or at the subjects' home.

Transcription and labeling

The collected definitions were first transcribed with Dragon Naturally Speaking 11.0 using a re-speaking method (the transcriber repeats the subject utterances into a microphone and the software performs speech-to-text transcription) and then manually corrected. As a drawback of the higher spontaneity typical of oral productions, subjects often produced long, continuous descriptions of stimuli, which therefore needed a delicate process of chunking into component features. Every

description was manually split ensuring that each concept's feature appeared on a separate line. For example, if a participant listed an adjective-noun property (*ha un collo lungo* “has a long neck”), the latter was divided into the features *lungo* “long” and *collo* “neck”, since adjective and noun correspond to different pieces of information regarding the concept. Some difficulties were encountered in determining whether certain properties could be regarded as light verb constructions with a verb carrying little semantic information (like *fare la doccia* “to take a shower” and *fare colazione* “to have breakfast”), and thus treated as a single feature, or as distinct properties (e.g., *andare in bicicletta*, literally “to go on bike”, was divided into the features *andare* “to ride” and *bicicletta* “bicycle”). The difficult cases were resolved by consensus between the authors.

Each feature was transformed into a normalized form: Nouns and adjectives were mapped to their singular and masculine forms, verbs to their infinitival forms; the infinitive passive form was used in the case where the concept requires a passive form. For instance, many participants said that *la carota è mangiata* “the carrot is eaten”. Since the carrot “undergoes” the action of being eaten, it was decided to use the infinitival passive form *essere_mangiato* “to_be_eaten” as the normalized feature. We maintained the original form only in cases in which the subjects produced a personal evaluation (e.g., *non mi piace* “I don't like it” for the concept *mela* “apple”).

The normalized features were translated into English and labeled with a semantic type. The BLIND feature type annotation scheme (reported in Appendix C) consists of 19 types that mark the semantic relation between the feature and the stimulus. A variety of annotation schemes have been proposed to capture relevant semantic aspects of the subject-generated features in previous norming studies. McRae et al. (2005) categorized the features with a modified version of the knowledge-type taxonomy later published in Wu and Barsalou (2009), and with the brain region taxonomy of Cree and McRae (2003). The former was also adopted and extended by Kremer and Baroni (2011). The STaR.sys scheme proposed by Lebani

and Pianta (2010) combines into a unique taxonomy insights from Cree and McRae (2003), Wu and Barsalou (2009), and lexical semantics (e.g., WordNet relations). Vinson and Vigliocco (2008) instead coded features into five categories referring to their sensory-motor and functional role. The same scheme was adopted by Montefinese et al. (2012) to code their Italian norms.

The BLIND scheme is inspired by Wu and Barsalou (2009) and Lebani and Pianta (2010), but it underspecifies certain semantic distinctions, thereby resulting into a smaller number of feature types (cf. 26 types in Lebani and Pianta's StaR.sys, 37 types in Wu & Barsalou, 2009, but 19 types in the BLIND scheme). The BLIND annotation scheme distinguishes among five macro-classes of feature types, each characterized by a set of more specific types:

1. **Taxonomical features**, including: the superordinate category of a concept (*mela* “apple” – *frutto* “fruit”: **isa** for “*is a*”); concepts at the same level in the taxonomic hierarchy, that is, coordinates or co-hyponyms (*cane* “dog” – *gatto* “cat”: **coo** for *coordinate*); concepts that have the stimulus concept as superordinate, that is, subordinates or examples of the stimulus (*nave* “ship” – *portaerei* “aircraft carrier”: **exa** for *example*); (approximate) synonyms (*afferrare* “to grab” – *prendere* “to take”: **syn** for *synonym*); concepts that are opposites of the stimulus concept on some conceptual scale, that is, antonyms (*libertà* “freedom” – *schiavitù* “slavery”: **ant** for *antonym*); and specific instances of a concept, expressed by proper nouns (*montagna* “mountain” – *Alpi* “Alps”: **ins**, for *instance*). The least frequently produced taxonomic types, **ant**, **coo**, **exa**, **ins** and **syn**, have been merged into the broader feature type **tax** in the coarser feature typology we use for the analysis below (given its prominence, **isa** has instead been maintained as a distinct type).
2. **Entity features** describe parts and qualities of the entity or event denoted by the stimulus concept (or larger things the concept is part of). Entity features thus include three “part-of” relations: meronyms, or parts of the concept (*cane*

“dog” – *zampa* “leg”: **mer** for *meronym*); holonyms, or larger things the concept constitutes a part of (*appartamento* “apartment” – *edificio* “building”): **hol** for *holonym*); and materials that the concept is made of (*spiaggia* “beach” – *sabbia* “sand”: **mad** for “*made of*”). These three types have been merged into the coarser feature **par** (*part*) for the analysis below. The other entity features pertain to qualities of the concept. All features corresponding to qualities that can be directly perceived, such as magnitude, shape, taste, texture, smell, sound and color, were grouped into “perceptual properties” (*canarino* “canary” – *giallo* “yellow”, *mela* “apple” – *rotondo* “round”, *ananas* “pineapple” – *dolce* “sweet”: **ppe** for “*property of perceptual type*”). All the qualities that were not perceptual were grouped into a class of “non-directly perceptual properties” (**pnp**, for “*property of non-perceptual type*”). The latter include abstract properties (*passione* “passion” – *irrazionale* “irrational”) or properties that refer to conditions, abilities or systemic traits of an entity, which cannot be apprehended only with direct perception and require some kind of inferential process (*carota* “carrot” – *nutriente* “nutritious”).

3. **Situation features** consider various aspects of the typical situations and contexts in which the concept is encountered (or takes place, if it is a verb), including events or abstract categories it is associated with, other participants in the event, and so on. Situation features thus include events (*nave* “ship” – *viaggiare* “to travel”: **eve** for *event*) and abstract categories associated with the stimulus (*duomo* “cathedral” – *fede* “faith”: **eab** for “*associated entity of abstract type*”), concrete objects that participate in the same events and situations with the stimulus concept (*martello* “hammer” – *chiodo* “nail”: **eco** for “*associated entity of concrete type*”), manners of performing an action associated with or denoted by the stimulus concept (*accarezzare* “to stroke” – *delicatamente* “gently”: **man** for *manner*), typical spaces (*nave* “ship” – *mare*

“sea”: **spa** for *space*) and temporal spans (*ciliegia* “cherry” – *maggio* “May”: **tim** for *time*) in which the concept is encountered or takes place.

4. **Introspective features**, expressing the subject’s evaluation of or affective/emotional stance towards the stimulus concept (*ciliegia* “cherry” – *mi piace* “I like it”: **eva** for *evaluation*).
5. **Quantity features**, expressing a quantity or amount related to the stimulus or to one of its properties (*motocicletta* “motorcycle” – *due* “two (wheels)”: **qua** for *quantity*).

As can be seen, there are various underspecified feature types. For instance, BLIND groups all perceptual qualities under the type **ppe**, without coding their sensory modality (cf. Lebani & Pianta, 2010; Vinson & Vigliocco, 2008). Moreover, differently from most annotation schemes (cf. Wu & Barsalou 2009), BLIND codes with **eve** all features expressing an event associated with the stimulus, independently of whether it represents the behavior of an animal (e.g., barking for a dog), the typical function of an object (e.g., transporting for a car), or simply some situation frequently involving an entity (e.g., cars and parking). The choice of underspecification in BLIND was motivated by the goal of optimizing the coders’ consistency in order to obtain reliably annotated features, while leaving to further research the analysis of more fine-grained feature types. The current coding allows researchers to select subsets of features that belong to a broad class (e.g., all perceptual properties) with good confidence in the quality and consistency of the annotation. The researchers can then perform a more granular (and possibly more controversial) annotation of this subset only, for their specific purposes (of course, we welcome contributions from other researchers that add a further layer of annotation to BLIND).

Two coders labeled the semantic relation types to reduce possible subjective errors. Problematic cases were discussed and resolved by consensus. To evaluate the inter-coding agreement, 100 randomly sampled concept-feature pairs were independently coded by the two annotators (of course, the coders did not

communicate during the agreement evaluation). Agreement between annotators is 0.76, with Cohen's kappa of 0.73, thereby adjusting the proportion of agreement for the chance agreement factor. A kappa value of 0 means that the obtained agreement is equal to chance agreement; a positive value means that the obtained agreement is higher than chance agreement, with a maximum value of 1. Although there is a lack of consensus on how to interpret kappa values, the 0.73 obtained for the BLIND annotation scheme is normally regarded as corresponding to substantial agreement (Artstein & Poesio, 2008).

Comparison with the Kremer and Baroni norms

Our subjects produced concept descriptions orally, and they were in general older than the undergraduates typically recruited for conceptual norming studies. One might wonder if these conditions resulted in data that are considerably different from those recorded in other conceptual norms, not only for blind, but also for sighted subjects. To explore the issue, we compared the concept-property pairs produced by our sighted subjects to those produced in written format by the younger Italian subjects of Kremer and Baroni (2011) (henceforth: KB). On the one hand, the shared language makes this comparison more straightforward than that with any other set of conceptual norms. On the other hand, KB showed that the data in their Italian norms are very similar to those in comparable German and English norms. Thus, if our sighted subject productions correlate with theirs, it is reasonable to conclude that they are also not significantly different from those collected in other related norming studies for English and other languages.

The different elicitation modality (or perhaps older age) does affect the quantity of properties produced: 10.5 (SD: 6.2) on average for our sighted subjects vs. 4.96 (SD 1.86) for those of KB. This figure is however not only affected by genuine differences in the properties, but also by different choices when segmenting and normalizing them.

To check if there are also significant differences in terms of the actual features that were produced, we looked at those 11 concepts that are shared across our collection and KB, and counted the proportion of KB concept-property pairs (featuring one of the shared concepts) that were also attested in our norms (these analyses were conducted on the `concept-feature-measures.txt` file described in Appendix D, and on the equivalent KB data set; for coherence with KB, we only considered pairs produced by at least 5 subjects). The resulting proportion is of 73/110 (66%) KB pairs that are also in the productions of our sighted subjects. Because of the differences in segmentation and normalization, it is difficult to draw a major conclusion about the non-overlapping part of the norms, and we can take 66% overlap as a lower bound on the actual cross-norms agreement (our blind and sighted data, for which, as we will show below, the overlap figure is comparable, were instead post-processed in identical ways). Further evidence of the qualitative similarity between our sighted norms and KB comes from the comparison of production frequencies across the 73 overlapping concept-property pairs. The correlation coefficient for the number of subjects that produced these pairs in KB vs. in our sighted set is at a very high (and highly significant) 0.84.

The BLIND resource

The BLIND conceptual norms are freely available under a Creative Commons Attribution-ShareAlike license (<http://creativecommons.org/licenses/by-sa/3.0/>). The users, besides of course being allowed to download and investigate the norms, are also allowed to modify and redistribute them, as long as they make the resulting resource available under the same kind of license, and they credit the original creators by citing this article.

Besides the concept description data and annotations described above, the norms contain numerous measures of concepts and features that might be useful to researchers, such as frequency and familiarity information, concept similarities in

terms of their production frequencies across features, and so on. The contents of the files included in the BLIND distribution are described in detail in Appendix D.

The norms can be downloaded from the website http://sesia.humnet.unipi.it/blind_data, that also offers a search interface to the concept and property data. This interface is meant to support qualitative explorations of the data, e.g., for the purpose of selecting experimental materials. Snapshots of the search interface and its output are presented in Figure 1.

BLIND - BLind Italian Norming Data

Feature-based descriptions of concepts produced by subjects in a property generation task are widely used in cognitive science to develop empirically grounded concept representations and to study systematic trends in such representations. BLIND is a publicly available collection of parallel semantic norms collected from a group of congenitally blind Italian subjects and highly comparable sighted subjects. The BLIND norms comprise descriptions of 50 nouns and 20 verbs.

Search by concept: banana Search

Search by PoS: Noun Verb Search

Search by concept class: abs_event Search

► Blind Sighted

Concept (It.)	Concept (Eng.)	Feature (It.)	Feature (Eng.)	Type	Count
banana	banana	frutto	fruit	isa:eco	22
banana	banana	essere_mangiato:essere_consumato	to_be_eaten	eve	10
banana	banana	allungato	oblong	ppe	8
banana	banana	africa	africa	spa	7
banana	banana	buccia	peel	mer	7
banana	banana	dolce	sweet	ppe:eco	5
banana	banana	essere_usato:essere_utilizzato	to_be_used	eve	5
banana	banana	ricurvo:curvo	curved	ppe	4
banana	banana	macedonia	fruit_salad	eco:spa:exa	4
banana	banana	contenere	to_contain	eve	4
banana	banana	paese_tropicale	tropical_country	spa	4

Figure 1. A snapshot of the online BLIND search interface (top) and partial output of the query for the concept *banana* (bottom).

A preliminary analysis of BLIND

The main purpose of this article is to introduce the publicly available BLIND database to the scientific community. However, in this section we also present an analysis of BLIND, to illustrate some general trends that bear on the theoretical debate on blind cognition, as well as a demonstration of the sort of investigation that our resource enables.

The global number of BLIND features (obtained with the segmentation process described above) is 19,087 for sighted subjects (4,508 distinct properties), and 17,062 for blind subjects (4,630 distinct properties). The higher number of features for sighted is due to the larger size of this group. In fact, there is no difference in the total number of features produced per-participant across the groups (average sighted: 734.1, SD: 285.3; average blind: 775.5, SD: 382.4; $t = -0.419$, $df = 38.305$, p -value = 0.6776). The average number of features produced per concept per participant is also not significantly different between the two groups (average sighted: 10.5, SD: 6.2; average blind: 11.9, SD: 7.4; $W = 1,429.768$, p -value = 0.3108). The two groups also behave alike for the number of stimuli for which no features were produced (61 in total across subjects for the blind, 71 for the sighted; $\chi^2 = 0.007941$, $df = 1$, p -value = 0.929). The proportion of concept-feature pairs produced (at least 5 times) by blind subjects that are also produced (with the same minimum threshold) by sighted subjects is of 74% (395/535), not much larger than the overlap reported above in the comparison with KB. The correlation between the production frequencies of overlapping concept-feature pairs in our sighted and blind groups is of 0.74. While high and highly significant, this correlation is lower than the one we reported above in the comparison of our sighted subjects with those of KB (0.84). Therefore, the “between-norm” correlation of sighted subjects in KB and BLIND is higher than the “within-norm” correlation of the two subject groups in BLIND (although, admittedly, the correlation between KB and sighted BLIND is based on 73 pairs, the one between blind and sighted BLIND groups on a larger set of 395 overlapping pairs, that might include more "difficult" items on which subjects might naturally show less agreement).

We compared the blind and sighted subjects in terms of all the global concept measures described in Table D2 (excluding those involving absolute counts of produced features, since the latter are trivially higher for the larger sighted group). The groups differ significantly in mean cue validity only (average blind: 0.01, SD:

0.009; average sighted: 0.009, SD: 0.01; Wilcoxon paired test $V = 1,548$, p -value = 0.04). The cue validity of a feature with respect to a concept is the conditional probability of the concept given the feature. Mean cue validity is the average cue validity across the features associated with a concept. Thus, the (small) difference between the two groups suggests that blind subjects produced features that are more diagnostic of the concepts they are produced for. Overall, however, the general lack of significant differences in the analysis of the proportional global concept measures suggests that blind and sighted subjects conducted the task in a very similar manner.

The absolute and relative frequencies of each annotated feature type in BLIND are reported in Table C1. The feature type distribution in the two subject groups can be explored with mosaic plots, which visualize the pattern of association among variables in contingency tables. Mosaic plots are extremely useful for multivariate categorical data analysis, providing a global visualization of the variable interactions which could hardly be obtained with other more standard methods, such as bar plots (Baayen, 2008), and were already adopted by Kremer and Baroni (2011) to analyze feature norms. In this paper, we limit our analysis to concrete and abstract nouns, but the same methodology can be applied to perform other investigations on the BLIND data. The mosaic plots in Figure 2, generated with the `vcd` package in the R statistical computing environment (Meyer, Zeileis, & Hornik, 2006), visualize the 2-way contingency table of the variables Subjects (rows), with values “blind” and “sighted”, and Feature Types (columns), with the values listed above and in Table C1 (for the current analysis, we used the broad feature types **tax** and **par**, respectively grouping taxonomic - except for **isa** - and “part-of” relations). The areas of the rectangles are proportional to the counts in the cells of the corresponding contingency tables. The widths of the rectangles in each row of the plots depict the proportions of features mapped to each of the feature types, for the respective subject group. The heights of the two rows reflect the overall proportion of features produced by the two groups.

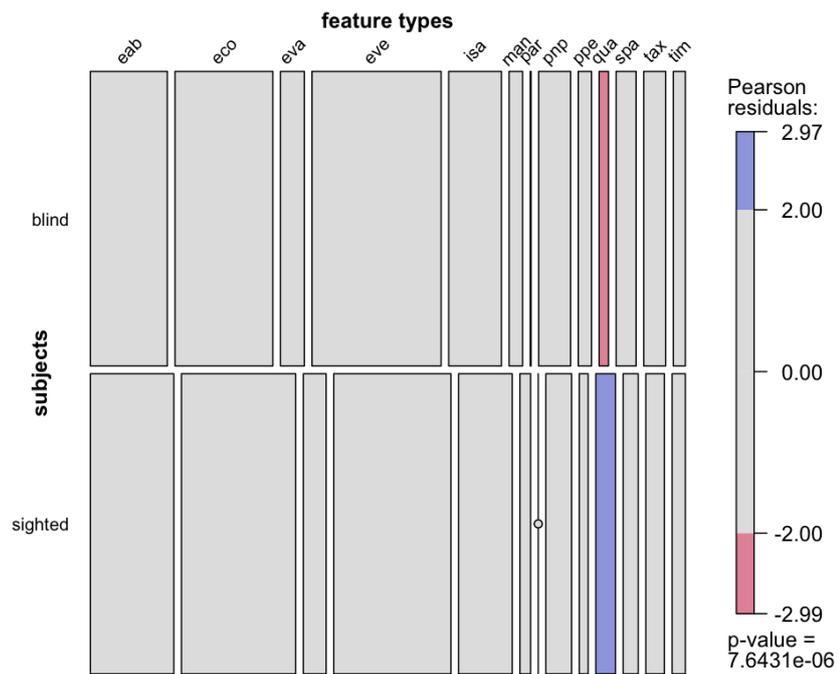
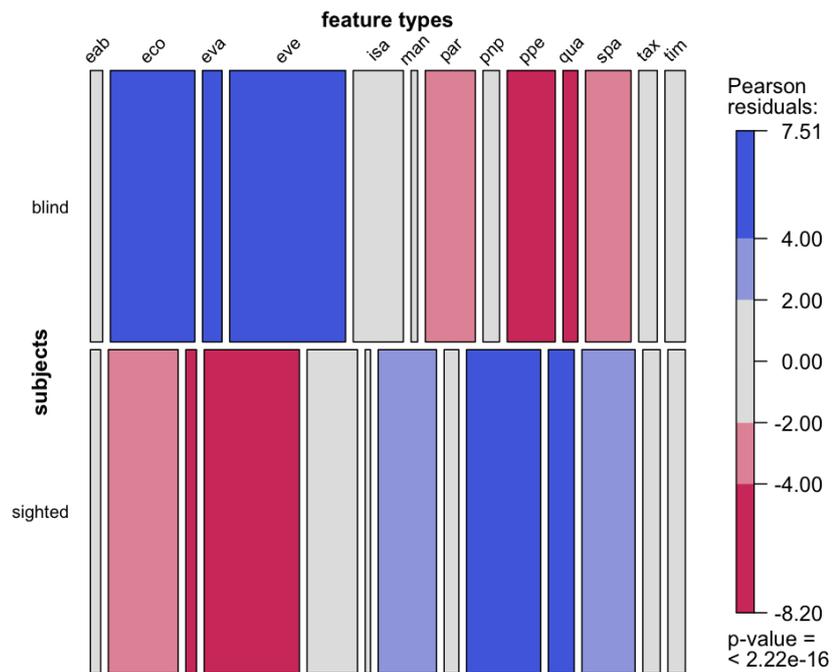


Figure 2. Overall frequency distribution of annotated feature types in sighted and blind subjects for concrete (top) and non-concrete nouns (bottom).

The color shades in the mosaic plots in Figure 2 code the degree of significance of the differences between the rectangles in a column according to a Pearson residual test, which compares the observed frequencies of features of a specific type to those expected given the feature distribution for both groups (Meyer et al., 2006). Blue rectangles correspond to significant positive deviances from the between-group expected distribution for a feature type. Red rectangles correspond to significant negative deviances from the between-group expected distribution. Color intensity is proportional to the size of the deviation from group-independence: dark colors represent large Pearson residuals (> 4), signaling stronger deviations, and light colors represent medium sized residuals ($2 < r < 4$), signaling weaker deviations. Grey represents a lack of significant differences between sighted and blind. The *p*-value of the Pearson residual test is reported at the bottom of the legend bar.

The top mosaic plot in Figure 2 represents the feature type distribution for the 40 concrete nouns. If we consider the areas of the rectangles, we notice that in both groups the lion's share is represented by situation features (**eva**, **eve**, **eco**, **man**, **tim**, and **spa**), followed by entity features (**par**, **ppe**, and **pnp**), and then by taxonomical ones (**tax** and **isa**). This essentially confirms previous analyses of property generation data (cf. Wu & Barsalou, 2009). On the other hand, the color shades also highlight important contrasts between the two groups. Sighted subjects differ from blinds for producing more perceptual (**ppe**) and quantity (**qua**) properties (dark blue rectangles in the bottom row), and to a lesser extent parts (**par**) and spatial (**spa**) features (light blue rectangles in the bottom row). Conversely, blind subjects show a robust preference for features corresponding to concrete objects (**eco**) or events (**eve**) associated with the stimulus, and for **eva** features expressing a subjective evaluation of the stimulus (dark blue rectangles in the bottom row). The two groups instead behave similarly with respect to the production of taxonomical (**tax** and **isa**) features, as well as for associated abstract entities (**eab**), temporal (**tim**) and manner (**man**) features, and for non-directly-perceptual properties (**pnp**). This is signaled by the grey

shade of the rectangles in the columns corresponding to these feature types.

Therefore, the BLIND data reveal that parts and directly perceptual properties are strongly underproduced by blind subjects to describe concrete nouns, when compared to sighted subjects.

Feature type production by the two groups for non-concrete nouns are represented in the bottom mosaic plot of Figure 2. Looking at the rectangle areas, we notice that parts (**par**), spatial features (**spa**) and perceptual properties (**ppe**) are strongly underproduced by both groups. This is obviously consistent with the fact that the stimuli express non-concrete concepts. Besides taxonomical features, the lion's share is now formed by associated events (**eve**) and objects (**eco**), in particular abstract ones (**eab**), as well as by subjective evaluation features (**eva**) and non-directly-perceptual properties (**pnp**). However, the most interesting fact is that, unlike with concrete nouns, no substantial differences emerge between blind and sighted subjects for ideal and emotion nouns (except for a mild distinction in quantity features, illustrated by the light blue and red colors of the relevant rectangles in the plot).

Conclusion

This article introduced BLIND, a collection of concept description norms elicited from Italian congenitally blind and comparable sighted subjects. The BLIND norms are documented in detail in Appendix D and freely available from http://sesia.humnet.unipi.it/blind_data. We also presented a preliminary analysis of the features in the norms, which is just a first step towards the new research avenues that open in front of us thanks to the BLIND data. For instance, blind and sighted subjects differ with respect to the *number* of directly perceptual properties they produce for concrete nouns (cf. Figure 2, **ppe** rectangle in top mosaic plot), but it is crucial to understand whether a difference exists also with respect to the *type* of perceptual properties the two groups produce (e.g., whether blind subjects produce fewer features related to vision than sighted subjects). To explore this issue, we are

currently carrying out a fine-grained annotation of the BLIND features with their dominant sensory modality, using existing modality norms for adjectives and nouns (Lynnott & Connell, 2009, 2012). Moreover, we intend to apply the annotation scheme proposed by Vinson and Vigliocco (2008) and analyze the distribution of sensory and functional features in BLIND, in order to understand whether blindness affects the way functional properties of objects are described. The difference between feature type productions we have reported concerns congenitally blind individuals. It is therefore natural to associate it with the total lack of visual experience by this group. We intend to extend BLIND with late blind subject data, in order to investigate whether comparable differences also exist between sighted and late blind subjects, thereby gaining new insights on the effect of visual deprivation on semantic memory.

As we have already said, feature norms by themselves do not commit to any particular form of underlying conceptual organization. Our first analysis of the BLIND data highlights important and significant differences between sighted and congenitally blind individuals and it is consistent with other experimental results, such as the ones reported in Connolly et al. (2007). However, further research is needed to understand which conclusions can be drawn from these differences with respect to the debate about embodied models of concepts, and the shape of semantic representations in the congenitally blind. Apart from any further speculation we can make, the BLIND data promise to become an important source of information on the relationship between concepts and sensory experiences, complementing and possibly enriching the evidence coming from the cognitive neurosciences.

References

- Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, *116*(3), 463-498. doi: 10.1037/a0016261
- Artstein, R., & Poesio M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics*, *34*(4), 555-596. doi: 10.1162/coli.07-034-R2
- Baayen, H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.
- Baroni, M., Barbu, E., Murphy, B., & Poesio, M. (2010). Strudel: A distributional semantic model based on properties and types. *Cognitive Science*, *34*(2):222–254. doi. 10.1111/j.1551-6709.2009.01068.x
- Baroni, M., Bernardini, S., Comastri, F., Piccioni, L., Volpi, A., Aston, G., & Mazzoleni, M. (2004). Introducing the “la Repubblica” corpus: A large, annotated, TEI(XML)-compliant corpus of newspaper Italian. In *Proceedings of LREC 2004*, Lisboa (1771-1774). Available from <http://www.lrec-conf.org/proceedings/lrec2004/pdf/247.pdf>
- Baroni, M., Bernardini, S., Ferraresi A., & Zanchetta, E. (2009). The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, *43*(3), 209-226. doi:10.1007/s10579-009-9081-4
- Baroni, M., & Lenci, A. (2008). Concepts and properties in word spaces. *Italian Journal of Linguistics*, *20*(1), 55-88. Available from: <http://poesix1.ilc.cnr.it/ijlwiki/doku.php>
- Baroni, M., & Lenci, A. (2010). Distributional Memory: A general framework for corpus-based semantics. *Computational Linguistics*, *36*(4), 673-721. doi: 10.1162/coli_a_00016
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577-660.

- Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London: Biological Sciences*, 358, 1177-1187. doi: 10.1098/rstb.2003.1319.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617-645. doi: 10.1146/annurev.psych.59.103006.093639
- Bedny, M., Caramazza, A., Pascual-Leone, A., & Saxe, R. (2012). Typical neural representations of action verbs develop without vision. *Cerebral Cortex*, 22(2), 286-293. doi: 10.1093/cercor/bhr081
- Cattaneo, Z., & Vecchi, T. (2011). *Blind vision. The neuroscience of visual impairment*. Cambridge, MA: The MIT Press.
- Chang, K. K., Mitchell, T., & Just, M. A. (2010). Quantitative modeling of the neural representation of objects: How semantic feature norms can account for fMRI activation. *Neuroimage*, 56(2), 716-727. doi: 10.1016/j.neuroimage.2010.04.271
- Connolly, 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(20), 8241-8246. doi: 10.1073/pnas.0702812104
- Cree, G., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, 132(2), 163-201. doi: 10.1037/0096-3445.132.2.163
- De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M. J., Voorspoels, W., & Storms, G. (2008). Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. *Behavior Research Methods*, 40(4), 1030-1048. doi: 10.3758/BRM.40.4.1030

- De Vega, M., Glenberg, A., & Graesser, A. C. (Eds.) (2008). *Symbols and embodiment: Debates on meaning and cognition*. Oxford, UK: Oxford University Press.
- Frassinelli, D., & Lenci, A. (2012). Concepts in context: Evidence from a feature-norming study. In *Proceedings of CogSci 2012*, Sapporo, Japan. Available from: <http://sesia.humnet.unipi.it/norms/>
- Gallese, V., & Lakoff, G. (2005). The brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3), 455-479. doi: 10.1080/02643290442000310
- Garrard, P., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2001). Prototypicality, distinctiveness, and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts. *Cognitive Neuropsychology*, 18(2), 125–174. doi:10.1080/02643290125857
- Kremer, G., & Baroni M. (2011). A set of semantic norms for German and Italian. *Behavior Research Methods*, 43(1), 97-119. doi: 10.3758/s13428-010-0028-x
- Landau, B., & Gleitman, L. R. (1985). *Language and experience. Evidence from the blind Child*. Cambridge, MA: Harvard University Press.
- Lebani, G. E., & Pianta E. (2010). A Feature Type Classification for Therapeutic Purposes: a preliminary evaluation with non-expert speakers. In *Proceedings of the Fourth Linguistic Annotation Workshop*, Uppsala, Sweden (157-161). Available from <http://aclweb.org/anthology-new/W/W10/W10-1823.pdf>
- Lynott, D. & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods*, 41(2), 558-564. doi: 10.3758/BRM.41.2.558
- Lynott, D. & Connell, L. (2012). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods*, 44(4). doi: 10.3758/s13428-012-0267-0

- Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-Specific Organization in the Human Brain Does Not Require Visual Experience. *Neuron*, *63*(3), 397-405. doi: 10.1016/j.neuron.2009.07.012
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology – Paris*, *102*, 59-70. doi: 10.1016/j.jphysparis.2008.03.004
- 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. doi: 10.1016/0022-0965(78)90082-6
- Marques, J. F. (2009). The Effect of Visual Deprivation on the Organization of Conceptual Knowledge. *Experimental Psychology*, *57*(2), 83-88. doi: 10.1027/1618-3169/a000011
- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547-559. doi: 10.3758/BF03192726
- Meyer, D., Zeileis, A., & Hornik, K. (2006). The strucplot framework: Visualizing multi-way contingency tables with vcd. *Journal of Statistical Software*, *17*(3), 1-48. doi: <http://www.jstatsoft.org/v17/i03>
- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2012). Semantic memory: A feature-based analysis and new norms for Italian. *Behavior Research Methods*, *44*(4). doi: 10.3758/s13428-012-0263-4
- Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge, MA: The MIT Press.
- Noppeney, U., Friston, K. J., & Price, C. J. (2003). Effects of visual deprivation on the organization of the semantic system. *Brain*, *126*(7), 1620-1627. doi: 10.1093/brain/awg152
- Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W.-H. C., Cohen, L., Guazzelli, M., & Haxby, J. V. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National*

- Academy of Sciences of the United States of America*, 10(15), 5658-5663. doi: 10.1073/pnas.0400707101
- Pietrini, P., Ptito, M., & Kupers, R. (2009). Blindness and consciousness: New light from the dark. In S. Laureys & G. Tononi (Eds.), *The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology* (pp. 360-374). Amsterdam, The Netherlands – Boston, MA: Elsevier Academic Press.
- Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22, 253-336.
- Ricciardi, E., Bonino, D., Sani, L., Vecchi, T., Guazzelli, M., Haxby, J. V., Fadiga, L., & Pietrini, P. (2009). Do we really need vision? How blind people “see” the actions of others. *Journal of Neuroscience*, 29, 9719-9724. doi: 0.1523/JNEUROSCI.0274-09.2009
- Riordan, B., & Jones, M. N. (2010). Redundancy in perceptual and linguistics experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science*, 3(2), 1-43. doi: 10.1111/j.1756-8765.2010.01111.x
- Ruts, W., De Deyne, S., Ameel, E., Vanpaemel, W., Verbeemen, T., & Storms, G. (2004). Dutch norm data for 13 semantic categories and 338 exemplars. *Behavior Research Methods*, 36(3), 506-515. doi: 10.3758/BF03195597
- Santos, A., Chaigneau, S. E., Simmons, W. K., & Barsalou, L. W. (2011). Property generation reflects word association and situated simulation. *Language and Cognition*, 3(1), 83-119. doi: 10.1515/langcog.2011.004
- Shepard, R. N., & Cooper, L. A. (1992). Representation of Colors in the Blind, Color-Blind, and Normally Sighted. *Psychological Science*, 3(2), 97-104. doi: 10.1111/j.1467-9280.1992.tb00006.x
- Steyvers, M. (2010). Combining feature norms and text data with topic models. *Acta Psychologica*, 133(3), 234-243. doi: 10.1016/j.actpsy.2009.10.010

- Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*, 422-488. doi: 10.1016/j.cogpsych.2003.09.001
- Vigliocco, G., & Vinson, D. P. (2007). Semantic representation. In G. M. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 195-215). Oxford, United Kingdom: Oxford University Press.
- Vinson, D. P., Vigliocco, G., Cappa, S., & Siri, S. (2003). The breakdown of semantic knowledge: Insights from a statistical model of meaning representation. *Brain and Language*, *86*, 347-365. doi: 10.1016/S0093-934X(03)00144-5
- Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*(1), 183-190. doi: 10.3758/BRM.40.1.183
- Wu, L. L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, *132*, 173-189. doi: 10.1016/j.actpsy.2009.02.002
- Zimler, J., & Keenan, J. M. (1983). Imagery in the congenitally blind: How visual are visual images? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(2), 269-282.

Appendix A
Concept stimuli

Table A1

Noun stimuli with their English translation (grouped by semantic class)

Concept class	Concepts
Artifactual location	<i>appartamento</i> “apartment”, <i>bar</i> “bar”, <i>duomo</i> “cathedral”, <i>negozio</i> “shop”, <i>ristorante</i> “restaurant”
Bird	<i>canarino</i> “canary”, <i>cigno</i> “swan”, <i>corvo</i> “crow”, <i>gabbiano</i> “seagull”, <i>pinguino</i> “penguin”
Emotion	<i>allegria</i> “cheerfulness”, <i>dolore</i> “pain”, <i>gelosia</i> “jealousy”, <i>passione</i> “passion”, <i>preoccupazione</i> “worry”
Fruit	<i>ananas</i> “pineapple”, <i>banana</i> “banana”, <i>ciliegia</i> “cherry”, <i>kiwi</i> “kiwi”, <i>mela</i> “apple”
Ideal	<i>amicizia</i> “friendship”, <i>democrazia</i> “democracy”, <i>giustizia</i> “justice”, <i>libertà</i> “freedom”, <i>religione</i> “religion”
Mammal	<i>cane</i> “dog”, <i>cavallo</i> “horse”, <i>gatto</i> “cat”, <i>giraffa</i> “giraffe”, <i>zebra</i> “zebra”
Natural location	<i>bosco</i> “woods”, <i>mare</i> “sea”, <i>montagna</i> “mountain”, <i>prato</i> “lawn”, <i>spiaggia</i> “beach”
Tool	<i>cacciavite</i> “screwdriver”, <i>coltello</i> “knife”, <i>matita</i> “pencil”, <i>martello</i> “hammer”, <i>pettine</i> “comb”
Vegetable	<i>carota</i> “carrot”, <i>lattuga</i> “lettuce”, <i>melanzana</i> “eggplant”, <i>patata</i> “potato”, <i>pomodoro</i> “tomato”
Vehicle	<i>aeroplano</i> “airplane”, <i>automobile</i> “car”, <i>motocicletta</i> “motorcycle”, <i>nave</i> “ship”, <i>treno</i> “train”

Table A2

Verb stimuli with their English translation (grouped by semantic class)

Concept class	Concepts
Abstract event	<i>credere</i> “to believe”, <i>dubitare</i> “to doubt”, <i>odiare</i> “to hate”, <i>pensare</i> “to think”, <i>temere</i> “to fear”
Auditory event	<i>ascoltare</i> “to listen”, <i>cantare</i> “to sing”, <i>gridare</i> “to shout”, <i>origliare</i> “to eavesdrop”, <i>udire</i> “to hear”
Tactile event	<i>accarezzare</i> “to stroke”, <i>afferrare</i> “to grab”, <i>massaggiare</i> “to massage”, <i>sfiorare</i> “to touch lightly”, <i>tastare</i> “to feel”
Visual event	<i>avvistare</i> “to spot”, <i>intravedere</i> “to glimpse”, <i>sbirciare</i> “to peep”, <i>scorgere</i> “to catch sight of”, <i>scrutare</i> “to peer at”

Appendix B

English translation of task instructions

Hi! You are going to hear a set of Italian words, and your task is to explain their meaning.

Please, remember these simple instructions:

1. do not hurry! Your task is to describe the word meaning as carefully as possible;
2. do not answer the first thing that comes to your mind! After listening to a word, think carefully about its meaning and those aspects that you regard as most important to describe it;
3. when explaining the word meaning, imagine that you are answering questions such as the following: what is it? what is it used for? how does it work? what are its parts? what is its shape? where can it be found? etc.;
4. describe each aspect of word meaning by using short sentences like these: *it is an animal, it is red, it has wings, etc.*

Here are some examples of descriptions:

- *trout*: it is a fish, it lives in rivers, it is good to eat, it can be fished, it has fins, it has gills, it has a silver color.

- *table*: it is a piece of furniture, it has usually four legs, it can be made of wood, it can be made of metal, it is used to put objects on it.

- *to eat*: it is an action, it is performed with the mouth, it is necessary for survival, it is pleasant, we use the fork, we use the knife, we can do it at a restaurant, we can do it at home.

Remember, in this task there is no right answer! You are absolutely free to explain as you wish what you believe to be the meaning of these words. You have one minute to describe each word: then you will hear a “beep” and you will hear the next word. If you think you are finished describing a word, you can move on to the next one by pressing the mouse left key. If you did not understand a word, you can listen to it again by pressing the mouse right key.

If the task is clear, you can now start a short trial session, after which you will be able to ask the experimenter for further explanations. If you want to listen to these instructions again, please press the mouse right key. If instead you are ready for the trial session, click the mouse left key.

Appendix C

BLIND feature type coding scheme

Table C1

Semantic feature types used to annotate the data from sighted (S) and blind (B) subjects, together with their absolute and relative occurrence frequencies

Feature type	Code	Examples	Group	No.	%
Hypernym	isa	<i>cat</i> -ANIMAL	S	2,103	11.02
			B	1,809	10.60
Coordinate	coo	<i>cat</i> -DOG	S	167	0.87
			B	175	1.02
Synonym	syn	<i>mountain</i> - MOUNT	S	29	0.15
			B	41	0.24
Antonym	ant	<i>hate</i> -LOVE	S	7	0.04
			B	16	0.09
Example_of	exa	<i>horse</i> -ROAN	S	415	2.17
			B	363	2.13
Instance	ins	<i>mountain</i> -ALPS	S	8	0.04
			B	8	0.05
Meronym	mer	<i>car</i> -WHEEL	S	1,175	6.16
			B	871	5.10
Holonym	hol	<i>seagull</i> -FLOCK	S	40	0.21
			B	21	0.12
Made_of	mad	<i>comb</i> -PLASTIC	S	236	1.24
			B	155	0.91
Perceptual_property	ppe	<i>cat</i> -BIG	S	1,988	10.41
			B	1,142	6.69
Non- directly_perceptual_property	pnp	<i>cat</i> -SOCIAL	S	633	3.32
			B	699	4.10

Entity_concrete	eco	<i>airplane</i> -PERSON	S	3,606	18.89
			B	3,383	19.83
Entity_abstract	eab	<i>dog</i> -FRIENDSHIP	S	1,027	5.38
			B	1,107	6.49
Event	eve	<i>bird</i> -FLY	S	3,723	19.50
			B	3,903	22.87
Subjective_evaluation	eva	<i>lion</i> -FEAR	S	478	2.50
			B	655	3.84
Space	spa	<i>zebra</i> -SAVANNA	S	1,643	8.61
			B	1,291	7.57
Time	tim	<i>tomato</i> -SUMMER	S	578	3.03
			B	568	3.33
Manner	man	<i>stroke</i> -GENTLY	S	388	2.03
			B	409	2.40
Quantity	qua	<i>car</i> -FOUR (wheels)	S	843	4.42
			B	446	2.61

Appendix D

Archived materials

The BLIND data files described in this appendix can be downloaded from http://sesia.humnet.unipi.it/blind_data. BLIND is released under a Creative Commons Attribution ShareAlike license (<http://creativecommons.org/licenses/by-sa/2.5/>).

The BLIND norms comprise the annotated production data (*production-data.txt*), separate measures for concepts (*concept-measures.txt*) and features (*feature-measures.txt*), measures for each concept's features (*concept-feature-measures.txt*), and concept pair similarities (*concepts-cosine-matrix.txt*). The measures are analogous to those reported by McRae et al. (2005) and Kremer and Baroni (2011), for comparability. All files are in simple text format, with variables arranged in columns separated by a tabulator space; the variable names are listed in the first line of each file.

Production data

The file *production-data.txt* contains annotated concept stimuli and subject responses as described in Table D1.

Table D1

Variables in the production data file

Variable name	Description
Subject	A unique numerical code identifying the subject.
Group	Whether subject is blind (b) or sighted (s).
ConceptIt	The stimulus concept name in Italian.
ConceptEn	English translation of the concept name.
POS	Part of speech of the concept (n if noun, v if verb).
Class	The class of the concept as detailed in the tables of appendix A above, with the following abbreviations: abs_event for abstract event, art_loc for artifactual location, aud_event for auditory event, nat_loc for natural location, tac_event for

	tactile event, vis_event for visual event.
Phrase	Transcription of the feature as produced by the subject (with minor editing, e.g., normalizing disfluencies).
FeatureIt	The feature in Italian in normalized format (see Section Transcription and labeling above for the normalization method).
FeatureEn	Translation of the normalized feature in English.
FeatureTypeGranular	The type of the relation between the feature and the concept, as detailed in Table C1.
FeatureTypeCoarse	Same as FeatureTypeGranular except that the types hol , mad and mer have been merged into par (part) and ant , coo , exa , ins and syn have been merged into tax (taxonomic excluding isa).
Polarity	Whether the subject asserted (<i>flies</i> ; value p) or negated the feature (<i>does not fly</i> ; value n)
FeatureOrder	Number recording the order in which the features were produced in the concept description by the subject.

Measures

All variables reported in the files to be documented next were extracted from a version of the production data from which we removed those rows containing a feature that was produced by a single blind or sighted subject for the corresponding concept (e.g., the row with concept *mare* “sea” and feature *sporco* “dirty” produced by sighted subject 1 was filtered out, because this was the only sighted subject who produced *dirty* as a feature of *sea*). Repeated feature productions by the same subject were also excluded from the computation of the relevant measures. Whenever a variable name has suffix Blind or Sighted, the corresponding measure was computed separately for the two subject groups.

Concept measures. The file *concept-measures.txt* contains measures pertaining to the concepts used in the norms. The variables in this file are described in Table D2.

Table D2

Variables in the concept measures file

Variable name	Description
ConceptEn	See Table D1.
ConceptIt	See Table D1.
POS	See Table D1.
Class	See Table D1.
LetterCount	Number of letters of the Italian concept name.
SyllableCount	Number of syllables of the Italian concept name.
FreqItWaC	Number of occurrences of the (Italian) concept name in the itWaC corpus (considering occurrences in both singular and plural form).
LogFreqItWaC	Natural logarithm of FreqItWaC.
FreqRepubblica	Number of occurrences of the concept name in the la Repubblica corpus.
LogFreqRepubblica	Natural logarithm of FreqRepubblica.
FreqWikipedia	Number of occurrences of the concept name in the Italian

	Wikipedia corpus.
LogFreqWikipedia	Natural logarithm of FreqWikipedia.
MeanFamiliarity(Blind Sighted)	Mean familiarity rating assigned by the subjects to the concept in a post-test survey. Subject rated the concepts on a 3-point scale, ranging from 3 (much experience with the object denoted by a noun or event/action denoted by a verb) to 1 (little or no experience).
SDFamiliarity(Blind Sighted)	Standard deviation of the familiarity ratings assigned to the concept.
FeaturesCount(Blind Sighted)	Number of distinct features produced for the concept.
DistinguishingFeaturesCount(Blind Sighted)	Number of distinguishing features produced for the concepts. A feature is distinguishing if it was produced for maximally two concepts.
DistinguishingFeaturesPercent(Blind Sighted)	Percentage of distinguishing features over overall feature count for the concept.
AverageDistinctivenessAcrossFeats(Blind Sighted)	Average distinctiveness across a concept features. The distinctiveness of a feature is the

	reciprocal of the number of concepts for which it was produced.
MeanCueValidity(Blind Sighted)	Average cue validity across the features of a concept. Cue validity is the conditional probability of a concept given a feature. It was calculated as the production frequency of a feature for a particular concept divided by the production frequency of that feature for all concepts.
IntercorrelatedFeatsCount(Blind Sighted)	Number of feature pairs of a concept for which features are intercorrelated, considering only those features appearing with at least three concepts. The correlation computation was based on the comparison of the feature production frequencies across the concepts. Feature pairs were considered correlated if they shared at least 6.5% of their variance (that is, the square of their Pearson correlation coefficient was at least 0.065).
IntercorrelatedFeatsPercent(Blind Sighted)	Percentage of intercorrelated feature pairs of a concept over

	the count of all the possible pairings of features of the concept.
IntercorrelationalDensity(Blind Sighted)	Sum of the percentage of shared variance across a concept intercorrelated feature pairs.

Feature measures. The file *feature-measures.txt* contains measures pertaining to the features produced by the subjects. The variables in this file are described in Table D3.

Table D3

Variables in the feature measures file

Variable name	Description
FeatureEn	English translation of the normalized feature name.
FeatureIt	Colon-delimited list of all Italian normalized feature names that were mapped to the same English translation, in order of production frequency (e.g., FeatureEn <i>belief</i> corresponds to FeatureIt <i>credenza:credo</i>).
FeatureTypeCoarse	Colon-delimited list of all coarse feature types (see Table D1) that were assigned to the feature, in order of production frequency.
FeatureTypeGranular	Colon-delimited list of all granular feature types that were assigned to the feature, in order of production frequency.
TotCount(Blind Sighted)	Overall production frequency of the feature

	across all concepts (if the feature was not produced by one of the two subject groups, corresponding TotCount is 0).
ConceptCount(Blind Sighted)	Number of distinct concepts the feature was produced for (if the feature was not produced by one of the two subject groups, corresponding ConceptCount is 0).
NegatedCount(Blind Sighted)	Number of times the feature was negated (if feature was not produced by one of the two groups, corresponding NegatedCount value is NA).
NegatedPercentage(Blind Sighted)	Percentage of times the feature was negated over all the times the feature was produced (if feature was not produced by one of the two groups, corresponding NegatedPercentage value is NA).
Distinguishing(Blind Sighted)	Whether feature is distinguishing (D) or not (ND) according to the criterion spelled out in Table D2 (if feature was not produced by one of the groups, corresponding Distinguishing value is NA).
Distinctiveness(Blind Sighted)	Distinctiveness score of the feature, calculated as described in Table D2 (NA if feature was not produced by the relevant group).

Concept-feature measures. The file *concept-feature-measures.txt* contains measures pertaining to the attested combinations of concepts and features. All the concept-related variables described in Table D2 are repeated in *concept-feature-measures.txt*. The following feature-related variables (described in Table D3) are repeated in *concept-feature-measures.txt*: FeatureEn, DistinguishingBlind,

DistinguishingSighted, DistinctivenessBlind, DistinctivenessSighted. The additional variables in *concept-feature-measures.txt* are described in Table D4.

Table D4

Additional variables in the concept-feature measures file

Variable name	Description
FeatureIt	Colon-delimited list of all Italian normalized feature names that were mapped to the same English translation, in order of production frequency (names and corresponding counts limited to the cases in which the feature was produced in response to the concept).
FeatureTypeCoarse	Colon-delimited list of all coarse feature types that were assigned to the feature, in order of production frequency (types and corresponding counts limited to the occurrences of the feature with the concept).
FeatureTypeGranular	Colon-delimited list of all granular feature types that were assigned to the feature, in order of production frequency (same restriction as for coarse types).
ConceptFeatureCount(Blind Sighted)	Number of subjects who produced the feature for the concept (0 if the feature was not produced in response to the concept by this subject group).
ConceptFeatureRank(Blind Sighted)	ConceptFeatureCount-based rank of the feature among those produced for the concept (NA if the feature was not produced

	in response to the concept by this subject group).
NegatedConceptFeature(Blind Sighted)	Whether at least two subjects negated the concept-feature relation (NA if the concept-feature pair was never produced by this subject group).

Concept similarities. The file *concepts-cosine-matrix.txt* contains, for all concept pairs, a score quantifying the similarity of their production frequency distributions across features. Production frequency distributions are kept separated for the two subject groups, so that concepts can be compared within or across groups. In particular, the concepts as represented by their feature frequency distributions pooled across blind subjects are suffixed by *-b*, and they are suffixed by *-s* when they are represented by their feature distributions pooled across sighted subjects (e.g., *dog-b* is the concept of dog described by blind subjects, *dog-s* is the same concept described by sighted subject). Similarity is measured by the cosine of the vectors representing two concepts on the basis of feature production frequencies. Values range from 0 (vectors are orthogonal, no similarity) to 1 (identical concepts).

Each line of *concepts-cosine-matrix.txt* after the first presents a concept followed by its cosine with all concepts (including itself), in fixed order (the first line contains the ordered list of concepts). For example, the first 3 values on the *dog-b* row are 0.10556, 0.13283 and 0.12657. Since, as reported on the first line, the first 3 concepts in the ordered list are *airplane-b*, *airplane-s* and *apartment-b*, this means that *dog-b* has a similarity of 0.10556 with *airplane-b*, of 0.13283 with *airplane-s* and of 0.12657 with *apartment-b*.