



Semantic similarity and associated abstractness norms for 630 French word pairs

Dounia Lakhzoum¹ · Marie Izaute¹ · Ludovic Ferrand¹

Accepted: 17 September 2020 / Published online: 1 October 2020
© The Psychonomic Society, Inc. 2020

Abstract

The representation of abstract concepts remains a challenge, justifying the need for further experimental investigation. To that end, we introduce a normative database for 630 semantically similar French word pairs and associated levels of abstractness for 1260 isolated words based on data from 900 subjects. The semantic similarity and abstractness norms were obtained in two studies using 7-point scales. The database is organised according to word-pair semantic similarity, abstractness, and associated lexical variables such as word length (in number of letters), word frequency, and other lexical variables to allow for matching of experimental material. The associated variables were obtained by cross-referencing our database with other known psycholinguistic databases including Lexique (New et al., 2004), the French Lexicon Project (Ferrand et al., 2010), Wordlex (Gimenes & New, 2016), and MEGALEX (Ferrand et al., 2018). We introduced sufficient diversity to allow researchers to select pairs with varying levels of semantic similarity and abstractness. In addition, it is possible to use these data as continuous or discrete variables. The full data are available in the [supplementary materials](#) as well as on OSF (<https://osf.io/qsd4v/>).

Keywords Semantic similarity norms · Concreteness · Normative ratings · Abstract concepts · French word pairs

Introduction

Conceptual representation has been the focus of much study and debate for decades in the fields of semantics and psycholinguistics. In contrast with the holistic and non-decompositional view held by Collins and Loftus (1975), at the heart of the debate now are two seemingly opposite accounts of semantic representation: the distributional and the embodied accounts of conceptual representation (Harris, 1954; Firth, 1957; see Lenci, 2008; Andrews, Vigliocco, & Vinson, 2009; Andrews, Frank, & Vigliocco, 2014; Bruni, Tran, & Baroni, 2014; Lenci, 2018 for reviews of the distributional account. See Glenberg, 1997; Barsalou, 1999; Zwaan, 2004; Meteyard, Cuadraro, Bahrami, & Vigliocco,

2012; Pulvermüller, 2013; Ostarek & Huettig, 2019 for reviews of the embodied account). Both these accounts consider that feature and property overlap play a major role in the processing of meaning (see Vigliocco & Vinson, 2007; Vigliocco, Meteyard, Andrews, & Kousta, 2009 for reviews). Indeed, there is much evidence of this from semantic priming studies, widely regarded as the gold standard for studying semantic representation in the mind and brain (e.g., Hutchison et al., 2013; Kim, Yap, & Goh, 2019). However, for both sides of the spectrum, the representation of abstract concepts remains a challenge, hence the need for a database of source material enabling us to further our understanding of abstract concept representation.

Accounts of semantic representation

Holistic view and spreading of activation

According to the holistic view of semantic representation, for every element of the world—be it an object, an event, property, etc.—there is an abstract and symbolic lexical equivalent that acts as a referent in the conceptual system of the mind (Fodor, Garrett, Walker, & Parkes, 1980; Berg & Levelt, 1990; Roelofs, 1997; Levelt, Roelofs, & Meyer, 1999). In this view of one-to-one mapping, each referent represents a single

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13428-020-01488-z>) contains supplementary material, which is available to authorized users.

- ✉ Dounia Lakhzoum
dounia.lakhzoum@uca.fr
- ✉ Ludovic Ferrand
ludovic.ferrand@uca.fr

¹ Université Clermont Auvergne, CNRS LAPSCO, 34 avenue Carnot TSA 60401, F-63001 Clermont-Ferrand Cedex 1, France

node in a semantic network, with nodes linked according to their semantic similarity. For instance, the concept *fire* would be represented by a single node linked to related concepts or properties, such as *red*, also represented by a single node. Collins and Loftus (1975) described the mechanisms of semantic processing based on their theory of the spreading of activation in a network, according to which a concept, when it is processed, activates the path between related nodes at a speed proportional to the strength of the link between them. The assumption of semantic similarity in the spreading of activation theory accounts for both the strength of the link between nodes and the ensuing dynamics of activation for related concepts. Given that in this holistic view, each property or feature of a concept is represented by a single node, it is a view which contrasts with the decompositional or featural view.

Featural view

According to the featural view of semantic representation, words can be decomposed into a set of defining features or properties reflecting the meaning of the concept to which they relate (Smith, Shoben, & Rips, 1974). For instance, the concept *fire* would be decomposed according to its defining features such as <is hot> and <is red>. As with the holistic view, at the core of the featural view is semantic similarity, but in this case it is measured by the number of features two concepts have in common (Plaut, 1995; McRae, de Sa, & Seidenberg, 1997; Cree, McRae, & McNorgan, 1999; Vigliocco, Vinson, Lewis, & Garrett, 2004; Kiefer & Pulvermüller, 2012). The more features they share, the more semantically similar they are. In recent years, two seemingly opposite accounts of this featural view have dominated the debate on the nature of semantic representation, namely the distributional and the embodiment accounts. They differ from each other in respect of the information used to represent meaning. While distributional semantics relies on symbolic and linguistic features, embodiment relies on perceptual and sensory-motor states.

According to models of distributional semantics, meaning is the result of the statistical distribution of words across written and spoken language (see Andrew, Frank, & Vigliocco, 2014; Lenci, 2018 for reviews of this account, see also Lund & Burgess, 1996; Landauer & Dumais, 1997; Griffiths, Steyvers, & Tenenbaum, 2007; Mandera, Keuleers, & Brysbaert, 2017). The meaning of words is therefore defined in relation to other words, depending on their shared symbolic and linguistic features. According to the distributional hypothesis, words occurring in similar contexts have similar meanings (Harris, 1954). This use of intralinguistic relationships was successfully implemented in computational models of semantics (e.g., Hoffman, McClelland, & Lambon Ralph, 2018). Motivation for using algorithms such as latent semantic analysis (LSA; Landauer & Dumais, 1997) is the notion that

meaning can be extracted by computing semantic similarities between concepts (Louwerse, 2008, 2011; Louwerse & Jeuniaux, 2008, 2010; Rogers & McClelland, 2004; Kintsch, McNamara, Dennis, & Landauer, 2007). In addition, the close performance between computational models and human behaviour suggests these models are able, to some extent, to mimic the extraction of semantic representation from language (see Andrews et al., 2009; Binder, Conant, Humphries, Fernandino, Simons, Aguilar, & Desai, 2016).

This view of distributional semantics using amodal linguistic symbols as a proxy for representing meaning has been under fire, particularly from researchers subscribing to the theory of embodiment, for its lack of grounding in perceptual and motor states.

The embodied account of semantic representation defines meaning as grounded in perceptual and motor states derived from an individual's sensory experience (Barsalou, 1999; Glenberg, 1997; Zwaan, 2004; Kiefer & Pulvermüller, 2012; Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012). For instance, Pulvermüller, Shtyrov, and Ilmoniemi (2005) used brain-imaging techniques to show that brain areas responsible for motor actions of the face and leg are activated when action words such as *kick* or *lick* are processed. Evidence like this struggles, however, to explain the grounding mechanisms for abstract concepts, where there are no physical and sensory features (see Borghi & Pecher, 2011; Borghi, Binkofski, Castelfranchi, Cimatti, Scorolli, & Tummolini, 2017, for reviews).

The dichotomy between abstract and concrete concepts is not clear-cut (Della Rosa et al., 2010). The most commonly invoked criterion is tangibility, with concrete concepts referring to tangible entities that are perceptible via the senses, whereas abstract concepts are intangible. According to the dual-coding theory (Paivio, Yuille, & Madigan, 1968), concrete concepts trigger processing based on two informational systems, one visual, the other verbal, whereas abstract concepts are processed only in the verbal system. The context availability theory (Schwanenflugel, Harnishfeger, Stowe, 1988) posits that while concrete concepts refer to a definite number of contexts, abstract concepts are connected to varied contexts. Although true, this distinction can be considered reductive and contributes to the view that abstract concepts are poor in terms of features. More recently, with the interest shown in abstract concepts by grounded cognition, new elements of definition have emerged, according to which abstract concepts refer to intangible features such as emotions, events, social contexts, and introspective states (e.g., Barsalou & Wiemer-Hastings, 2005; Harpaiter et al., 2018; see Borghi et al., 2017 for a review). This latter definition reflects a new interest in their grounding mechanisms and semantic representation.

The tangibility criterion is best represented by the concreteness variable defining the distinction between concrete and

abstract concepts based on the dual coding and context availability theories. It plays a key role in psycholinguistic research, as well as providing an explanation for many phenomena, such as hemispheric lateralisation in the processing of concrete and abstract concepts (Oliveira, Perea, Ladera, & Gamito, 2013), or ease of retrieval of concrete words compared to abstract ones (Mate, Allen, & Baques, 2012; Nishiyama, 2013).

The importance of the concreteness variable is further borne out by the development of several widely used databases containing concreteness rating norms (Coltheart, 1981) and, more recently, 40,000 words in English (Brysbaert, Warriner, & Kuperman, 2014) and 1659 words in French (Bonin et al., 2018).

Abstract concept representation

The embodied account has yet to propose a unified theory for the representation of abstract concepts such as *justice* or *freedom* which do not refer to direct perceptual features or sensory-motor states (Dove, 2009, 2011, 2014; Machery, 2016; see Pecher, 2018 for a review). However, several hypotheses, ranging from strongly to weakly embodied, have been put forward as explanations for the grounding mechanisms of abstract concepts. The strong embodiment assumptions make no allowance for multiple representations and consider abstract concepts to be as grounded and reliant on sensory-motor systems as concrete concepts are (e.g., Glenberg & Kaschak, 2002; see Borghi et al., 2017 for a review). For instance, according to the conceptual metaphor theory, abstract concepts are grounded through image schemas corresponding to mental representations (e.g., Lakoff & Johnson, 1980; Gallese & Lakoff, 2005). Several studies have shown that abstract concepts of *valence* and *power* are grounded in two-dimensional spatial schema with the higher point of a vertical vector representing positions of power while the left-hand side of a horizontal vector represents negative concepts (see Pecher, 2018 for a review). However, the need for one-to-one mapping between abstract concepts and concrete metaphors means there are limits to the availability of such metaphors for every type of abstract concept.

At the other end of the spectrum, according to weak embodiment assumptions, abstract concepts are grounded via multiple representations of meaning with the involvement of both sensorimotor and linguistic processing. These grounding mechanisms place a greater emphasis on the context in which abstract concepts are used (e.g., Barsalou, 1999, 2003; Wiemer-Hastings & Xu, 2005). Several studies have shown that abstract concepts activate social and introspective aspects of situations (Barsalou & Wiemer-Hastings, 2005), emotional features (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011; Lenci, Lebani, & Passaro, 2018), information about

events, and thematic roles (Ferretti, McRae, & Hatherell, 2001), and, more generally, linguistic information acting as a shortcut to conceptual simulation (Barsalou, Santos, Simmons, & Wilson, 2008). Such assumptions have the advantage of being sufficiently general to apply to a variety of abstract concepts.

Whether seen from the distributional or embodied end of the spectrum, all accounts agree on the importance of relationships between concepts for the organisation of semantic knowledge. Two kinds of relationships have been widely investigated: semantic similarities (theft-burglar) and verbal association (theft-prison), and much effort has gone into creating databases of material to use in semantic priming studies regarded as the gold standard for studying how semantic knowledge is organised (see Hutchinson, Balota, Cortese, & Watson, 2008; Hutchison et al., 2013; Pulvermüller, 2013; Mandra, Keuleers, & Brysbaert, 2017).

Semantic priming and semantic similarity for concrete and abstract concepts

Semantic priming In a semantic priming study, participants are presented with a prime word followed by a target word (Meyer & Schvaneveldt, 1971). The relationship between the two words is one of either semantic similarity, where the two words belong to the same superordinate category (e.g., prime: *eagle*; target: *owl*), or verbal association, where the two words are frequently found together across spoken and written language (e.g., prime: *fireman*; target: *truck*; McNamara, 1992; Plaut, 1995). In a lexical decision task, participants make a decision on the target by indicating whether or not it is a word. The semantic priming effect refers to the robust result, which has been replicated hundreds of times, showing that participants respond faster for related primes and targets compared to unrelated ones (Hutchison et al., 2008, 2013). This phenomenon has been widely studied as it provides considerable insight into the organisation and mechanisms of semantic knowledge. In actual fact, each theoretical view discussed above can account for this priming effect. According to the holistic view (Fodor et al., 1980; Berg & Levelt, 1990; Roelofs, 1997), the priming effect is the result of spreading activation from the prime to the target along strongly linked nodes, whereas according to the distributional, embodied, and hybrid accounts, it results from the activation of features shared between the prime and the target (Mahon & Caramazza, 2008; Dove, 2009; Andrews, Frank, & Vigliocco, 2014; Carota, Kriegeskorte, Nili, & Pulvermüller, 2017). Where these last accounts differ, however, is in the nature of the features. The distributional account suggests the priming effect results from the activation of linguistic features, the embodied account that it results from the activation of shared sensorimotor states, and the hybrid account that both

linguistic and perceptual features are responsible for this phenomenon.

Despite the robustness of the semantic priming effect with concrete concepts, with abstract concepts results have been inconsistent. Crutch (2005; Crutch, Conell, & Warrington, 2009; Crutch and Warrington, 2010) showed that while concrete concepts are organised according to semantic similarity, abstract concepts are organised according to verbal association. Several studies tried to replicate these results but revealed discrepancies (e.g., Hamilton & Coslett, 2008; Duñabeitia, Avilés, Afonso, Scheepers, & Carreiras, 2009; Geng & Schnur, 2015). Indeed, these studies have attempted to replicate the results according to which concrete and abstract concepts have different dependencies upon semantic similarity and associative strength. They have however failed to find any such difference in the organisation of concrete and abstract concepts. A more recent study found that both semantic similarity and verbal association elicited a priming effect for concrete concepts, whereas for abstract concepts it was found only with verbal association (Ferré, Guasch, García-Chico, & Sánchez-Casas, 2015). Crutch and Jackson (2011) suggested the relationship between concreteness and association type could explain these disparities. They presented evidence based on data from healthy and neuropsychological patients showing that when presented with triplets of low, middle, and high levels of concreteness, the effect of semantic similarity increased with concreteness, while the effect of verbal association decreased with concreteness. Furthermore, they suggested that concreteness be used as a graded variable rather than a binary one, especially when studying its effect on the organisation of semantic memory. Accordingly, this calls for a shift in the way abstract concepts are studied, to place more emphasis on the type and associated level of concreteness for selected abstract concepts. Two different procedures are used to generate material for semantic similarity and priming studies: feature generation tasks and semantic similarity ratings.

Semantic similarity: feature generation and semantic pairs In a feature generation task, participants are given a list of words for which they are required to provide a list of features defining each word. The procedure provides measures of semantic similarity by comparing the feature overlap between two words. The more features two words have in common, the more similar they are (McRae, Cree, Seidenberg, & McNorgan, 2005; McRae, de Sa, & Seidenberg, 1997; Sánchez-Casas, Ferré, García-Albea, & Guasch, 2006; Vigliocco, Vinson, Lewis, & Garrett, 2004; Vinson & Vigliocco, 2008). However, it is a procedure which is highly time-consuming and which has limitations (see McRae et al., 2005 for a discussion of these limitations). For instance, in feature naming, participants may provide only a linguistic approximation of conceptual content. It is fair to assume, therefore, that some parts of the concepts would be lost in

verbalisation. This criticism appears to be particularly relevant in the case of abstract concepts which may themselves be decomposed into abstract features. Indeed, many authors have suggested that, compared to concrete concepts, abstract concepts appear to be semantically impoverished, with their representation requiring associations with other concepts or grounding simulations in introspective and social states (Barsalou et al., 2008; Borghi, Scorolli, Caligiore, Baldassare, & Tummolini, 2013; Borghi, Barca, Binkofski, Castelfranchi, Pezzulo, & Tummolini, 2019, see also Recchia & Jones, 2012).

On the other hand, Wiemer-Hastings and Xu (2005) suggested that this apparent paucity of features for abstract concepts is due mainly to the instructions given to participants during a feature generation task. In the original method, Wiemer-Hastings and Xu (2005) asked participants only to generate features defining the concept, whereas later they instructed them to provide context features. The results showed that the difference between abstract and concrete concepts in terms of semantic richness disappeared when participants were encouraged to provide context features. By using the same method of property listing as Wiemer-Hastings and Xu (2005), Harpainter, Trumpp, and Kiefer (2018) gathered properties for close to 300 abstract concepts. By doing so, they further demonstrated the richness and heterogeneity of abstract concepts, showing that they can elicit affective, introspective, social, and sensory-motor properties. This heterogeneity of abstract concepts was further investigated by Villani, Lugli, Liuzza, and Borghi (2019) who evaluated more than 400 abstract concepts on 15 dimensions. Their results provided further support for a multiple representation view of abstract concepts.

In addition, Bolognesi, Pilgram, and van den Heerik (2017) adapted Wu and Barsalou's taxonomy (2009) to include 20 feature categories belonging to four main dimensions (concept properties, situation properties, introspections, and taxonomic properties) that must be distinguished to convey the full semantic richness of concepts. Recchia and Jones (2012) were not, however, able to determine whether such distinctions in respect of semantic features could benefit abstract concept representation. They invoked the shallowness of lexical decision tasks in semantic processing. Consequently, future studies will need to reach conclusions on the role of feature categories for abstract concept representation.

Another, less costly, way of creating material for semantic representation studies is to generate semantically similar word pairs. This option relies on a similarity-rating task where participants are presented with pairs of words formed by the researcher with a view to obtaining concepts either belonging to the same category or being similar in meaning (e.g., *truck-car*; Ferrand & New, 2003; Perea & Rosa, 2002). Participants must rate the semantic similarity of the pairs on a scale (Ferrand & New, 2003; Sánchez-Casas et al., 2006). Studies

have shown that the pairs rated as being highly similar produced a strong priming effect (e.g., McRae & Boisvert, 1998; Plaut & Booth, 2000; Hutchison, 2003; Andrews, Lo, & Xia, 2017). In addition, studies have shown a strong correlation between the measures from similarity-rating tasks and feature generation, ensuring the legitimacy of this latter technique (e.g., McRae, de Sa, & Seidenberg, 1997). More recently, Maki, Krinsky, and Muñoz (2006) used a semantic rating task to show that ratings were a good predictor of feature overlap for existing semantic feature norms.

Normative databases for semantic similarity

Given the importance of carefully crafted material for studying semantic representation, much effort has been directed towards building normative databases to provide the research community with the material it needs. The most commonly found data sets gather English feature norms. McRae and collaborators (2005), for instance, provides feature norms for 541 living and non-living concepts. Subsequently, Buchanan, Holmes, Teasley, and Hutchison (2013) built a searchable web portal based on the work of McRae and collaborators (2005), facilitating the search for experimental stimuli in their data set. Buchanan, Valentine, and Maxwell (2019) expanded previous databases and provided features for more than 4000 words. Vinson and Vigliocco (2008) provided an interesting data set based on concrete object nouns and verb events that allow semantic representation to be studied beyond the usual focus on concrete concepts. Devereux, Tyler, Geertzen, and Randall (2014) built on McRae and colleagues' work by adding features produced by at least two participants compared to McRae and collaborator's (2005) five-feature threshold for inclusion. In other languages, De Deyne and Storms (2008) and De Deyne et al. (2008) collected normative features among Dutch participants. Lebani, Bondielli, and Lenci (2015) collected thematic role features to study the semantic content of Italian verbs. Also in Italian, Lenci, Baroni, Cazzolli, and Marotta (2013) collected semantic features from congenitally blind and sighted participants, making it possible to study the role of perceptual information in concept processing. Kremer and Baroni (2011) collected properties and semantic relation types for German and Italian. More recently, Vivas, Vivas, Comesaña, Coni, and Vorano (2017) published the first Spanish semantic feature production norms for living and non-living concepts.

Researchers have used similarity-rating tasks to a lesser extent to produce such norms. Buchanan and collaborators (2013) compiled an English data set comprising 1808 words paired according to semantic similarity. In Spanish, Moldovan, Ferré, Demestre, and Sánchez-Casas (2015) collected normative ratings for 185 Spanish noun triplets with variation of semantic distance within each triplet. However, much of the effort in developing databases has been focused

on concrete concepts. To the best of our knowledge, the present work offers the first database of semantically similar abstract word pairs in French.

The present study: Semantic similarity norms for abstract words

The present work introduces a data set comprising semantic similarity ratings for abstract word pairs obtained from French participants. We have added a measure of the concreteness of each word from each pair to allow for the selection of abstract concepts in line with Crutch and Jackson's (2011) suggestion that there is a relationship between graded levels of concreteness and semantic organisation. To provide a data set of experimental stimuli according to the significant lexical variables and lexical latencies previously discussed, we have combined our list of words with existing databases such as the French Lexicon Project (FLP, Ferrand et al., 2010), Lexique (New et al., 2001, 2004, 2007), MEGALEX (Ferrand et al., 2018), and Wordlex (Gimenes & New, 2016).

Method

Participants

Both the similarity- and concreteness-rating tasks were presented as online questionnaires. Participants for the two studies were all French native speakers and between 18 and 45 years old. We collected data from 373 participants (334 women; $M_{\text{age}} = 26.43$; $SD = 8.34$) for the similarity-rating task, and 529 (486 women; $M_{\text{age}} = 29.7$; $SD = 9.03$) for the concreteness-rating task. Participants volunteered in response to an announcement posted on Facebook group walls, and no compensation was paid. Participants took part in only one of the tasks in an attempt to ensure their ratings were not influenced by previous exposure to the items which are common to both tasks. Both studies obtained the approval of the Université Clermont Auvergne Research Ethics Committee.

Stimuli

To have some guarantee of the level of abstractness¹ of our material before collecting our own ratings, we selected 1020 words having a low level of concreteness (range between 100 and 600) from Coltheart's (1981) concreteness norms. We then translated the selected words into French following a back-translation procedure (Sperber, Devellis, & Boehlecke,

¹ Both "abstractness" and "concreteness" words are used throughout this paper. The use of these notions is not arbitrary. Abstractness is a central notion to the present work, as it aimed at introducing abstract stimuli, whereas concreteness is used to refer to other studies that introduce or deal with the concreteness variable.

1994), following which 174 words were excluded. We also added the material from Ferrand (2001) comprising 260 French abstract words.

Based on our linguistic intuition, we then formed semantically similar pairs (e.g., joie-bonheur; [*joy-happiness*]). To the best of our ability (see below), we ensured that the semantic pairs were non-associates (according to McRae & Boisvert, 1998), and were not linked by either a super/supra-ordinate, part/whole, or antonym relationship. The material was then divided into six lists of pairs, and 30% of fillers (unrelated pairs, e.g., défaut-frisson; [*flaw-chill*]) were added per list. So that the participants would be sensitive to the abstractness of the pairs, we also added concrete words from Ferrand and Alario (1998) and formed semantic pairs. Accordingly, we were able to form 628 semantically related pairs (460 noun pairs, 99 adjective pairs, and 69 verb pairs). Both prime and target words had the same grammatical status within each semantically similar pair. To ensure the pairs were semantically similar and not associated, we translated the target words back into English and checked for forward strength in the Small World of Words database² (SWOW, De Deyne, Navarro, Perfors, Brysbaert, & Storms, 2019). We identified all pairs for which the prime and target presented a forward associative strength of higher than 10%. Seventy pairs were identified as both associated and semantically similar (e.g., anxiety-fear). We kept them in the main database with the possibility to filter them out. In addition, we created a secondary database containing only the semantically similar and associated word pairs. As suggested by De Deyne et al. (2019), association data are not to be discarded and provide a strong indication of meaning similarity.

For the concreteness-rating task, the pairs were separated, and the lists of individual words were presented in another experiment. Given the added material from Ferrand and Alario (1998), participants were presented with stimuli ranging from abstract to concrete, thereby ensuring their sensitivity to the task and avoiding learned response patterns.

Procedure

The stimuli (fillers included) were randomly divided into 6 lists of word pairs and 10 lists of isolated words, respectively, for the similarity-rating and concreteness-rating tasks. The motivation for dividing the pairs into different lists was two-fold. Firstly, we wanted to keep the experiment concise so as to not overwhelm participants. Secondly, some words appear several times in different pairs, which is why we used semi-randomization to ensure that participants never saw pairs with the same words. The pairs and words were presented one by one on the screen in a randomised order. The experiment was

² We used the Small World of Words database because there are no databases large enough in French.

conducted online using the Qualtrics software (2020). The design of the interface for this experiment allowed participants to complete the task on either a computer or smartphone.

Once they had given their consent and registered their demographic information, participants were randomly assigned to one of the lists. Their task was to judge the similarity between the two words presented for the similarity ratings and whether the words were more abstract or concrete for the concreteness ratings. Both tasks used a 7-point Likert-like scale ranging from 1 = “not at all similar” (“*pas du tout similaires*” in French) to 7 = “totally similar” (“*tout à fait similaires*”) for the similarity-rating tasks and from 1 = “very abstract” (“*très abstrait*”) to 7 = “very concrete” (“*très concret*”) for the concreteness-rating task (see [supplementary material](#) for the specific instructions). The words appeared one by one on the screen and were replaced as soon as participants had rated them. They were presented in the middle of the screen in Arial 12 font against a white background. We provided examples of items and their possible ratings in the instructions. No training was given before the tasks started. Both studies were self-paced, with no time limit for either the stimulus presentation (word pair or isolated pair) or participant’s answer. Both tasks took about 12 minutes to complete.

Results

We first computed general statistics for the entire data set. The general statistics collected for the semantic similarity and concreteness variable are shown in Table 1³. Tables 2 and 3 provide the means for associated lexical variables computed by crossing our data set with the Lexique (New et al., 2004), FLP (Ferrand et al., 2010), MEGALEX (Ferrand et al., 2018), and Wordlex (Gimenes & New, 2016) databases.

It is apparent from the general statistics in Table 1 that the semantic similarity ratings range from 1.13 to 6.93 on a 7-point scale. This shows participants used the full range of the scale, but also reflects the diversity of the word pairs in terms of semantic similarity. Separating very similar ($M = 5.13$; $SD = 0.41$) and less similar ($M = 3.67$; $SD = 0.59$) pairs based on the median revealed a significant effect of semantic similarity [$t(300) = 35.78$, $p < 0.001$, $d = 2.06$]. This effect is particularly large, given that Cohen’s d suggests the difference is greater than two standard deviations. This will allow for the use of semantic similarity as either a continuous or categorical variable for researchers who would wish to study the effect of variation in semantic similarity. Concerning the

³ The semantic similarity variable reported in Table 1 corresponds to the mean similarity ratings for word pairs. The concreteness variable corresponds to the mean concreteness for the prime word and target word separately, as ratings were obtained on individual words for concreteness and on pairs of words for semantic similarity.

Table 1. Semantic similarity for word pairs and associated concreteness for prime and target words

	Pair semantic similarity	Prime word concreteness	Target word concreteness
Mean	4.43	4.41	4.40
SD	1.63	1.48	1.54
Min	1.13	1.63	1.51
Max	6.93	6.92	6.95
Median	4.65	4.43	4.47
Range	5.80	5.29	5.44
Skewness	-0.52	-0.31	-0.35
Q1	3.37	3.44	3.38
Q3	5.60	5.40	5.46

concreteness variable, the means for prime and target are very close to one another, showing a good concreteness match within each pair (mean prime concreteness = 4.41; mean target concreteness = 4.40). A paired-samples *t* test showed no significant difference between the mean concreteness ratings for prime and target words [$t(628) = 0.27, p = 0.80$ ns]. This close match is further demonstrated in the correlation we computed between prime and target words with a strong and highly significant correlation [$r = 0.87, t(628) = 44.50, p < 0.001$].

Tables 2 and 3 display the lexical characteristics for the primes and targets composing our word pairs. The statistics presented in Tables 2 and 3 were obtained by cross-referencing our data set with Lexique (New et al., 2001, 2004, 2007), the French Lexicon Project (Ferrand et al., 2010), Wordlex (Gimenes & New, 2016), and MEGALEX (Ferrand et al., 2018). Movie subtitle frequency corresponds to the *freqfilms2* variable from Lexique and refers to word frequency based on movie subtitles. The

other frequencies were computed from books (Lexique: New et al., 2004), blog posts, Twitter, and newspapers (Wordlex: Gimenes & New, 2016).

We also computed correlations between semantic similarity for the pair and lexical variables as well as concreteness levels for the prime and target respectively. Such correlations were all non-significant except for the correlation between semantic similarity and concreteness. Indeed, the concreteness level of the prime and target was negatively and moderately correlated with the semantic similarity of the pair, respectively ($R_{\text{prime_concreteness}} = -0.26; R_{\text{target_concreteness}} = -0.28, p < 0.001$), suggesting that the higher the semantic similarity, the lower the level of concreteness. However, the mean concreteness is not as different for highly similar pairs ($M_{\text{concreteness}} = 4.14; SD = 1.59$) as for less similar pairs ($M_{\text{concreteness}} = 4.78; SD = 1.39$). This means researchers using the present database will be able to study phenomena of semantic similarity and their relationship with graded levels of concreteness without having to worry that the concreteness variable and the lexical variables might act as confounding variables.

In addition, we computed correlations between the concreteness variable and other lexical variables. It is clear from Table 4 that the concreteness variable shows a negative correlation to frequencies based on blog posts and Twitter. Such correlations are rather weak ($r = -0.10$), however, and should not be cause for concern as regards potential confounding variables. The concreteness variable is also moderately and negatively correlated with the number of letters and orthographic similarity, but positively correlated with the number of orthographic neighbours. All lexical variables are significantly intercorrelated, a result which replicates previous findings from the psycholinguistic norms literature. Indeed, upon comparing the correlations shown in Table 4 with those

Table 2. Descriptive and behavioural data for target words

	Min	Max	M	SD
Movie subtitle frequency	0.00	2751.99	56.21	186.30
Book frequency	0.00	4696.15	64.51	286.62
Blog frequency	0.06	3095.93	84.85	279.23
Twitter frequency	0.00	4070.02	77.11	309.32
Newspaper frequency	0.00	3782.26	78.54	295.09
N-letters	3.00	13.00	6.88	2.03
N-orthographic neighbours	0.00	23.00	2.52	3.73
OLD20	1.00	5.40	2.06	0.62
Reaction times (FLP)	515.54	915.85	653.09	62.71
Reaction times (MEGALEX)	473.48	747.84	549.10	43.01

Note. OLD20: orthographic Levenshtein distance (Yarkoni, Balota, & Yap, 2008); Movie subtitle frequency (from Lexique: New et al., 2004)

Table 3. Descriptive and behavioural data for prime words

	min	max	M	SD
Movie subtitle frequency	0.00	986.59	31.02	81.43
Book frequency	0.00	835.47	35.36	81.71
Blog frequency	0.15	2866.60	53.28	150.60
Twitter frequency	0.07	4070.02	46.79	189.88
Newspaper frequency	0.00	1556.37	46.37	118.08
N-letters	3.00	14.00	6.99	2.00
N-orthographic neighbours	0.00	20.00	2.43	3.58
OLD20	1.00	4.55	2.10	0.60
Reaction times (FLP)	515.54	1005.15	657.94	67.18
Reaction times (MEGALEX)	473.92	738.03	553.11	44.78

Note. OLD20: orthographic Levenshtein distance (Yarkoni, Balota, & Yap, 2008); Movie subtitle frequency (from Lexique: New et al., 2004)

Table 4. Correlation matrix between concreteness levels and lexical variables with significance levels

	Concreteness	Movie Subtitles Frequency	Book Frequency	Blog Frequency	Twitter Frequency	Newspaper Frequency	N-letters	N-orthographic neighbours	OLD20
Concreteness	1								
Movie Subtitles Frequency	-0.03	1							
Book Frequency	-0.003	0.78***	1						
Blog Frequency	-0.10**	0.73***	0.73***	1					
Twitter Frequency	-0.10**	0.63***	0.45***	0.87***	1				
Newspaper Frequency	-0.06	0.68***	0.78***	0.89***	0.66***	1			
N-letters	-0.24***	-0.23***	-0.20***	-0.23***	-0.22***	-0.19***	1		
N-orthographic neighbours	0.18***	0.24***	0.24***	0.23***	0.24***	0.19***	-0.59***	1	
OLD20	-0.21***	-0.19***	-0.17***	-0.18***	-0.18***	-0.16***	0.79***	-0.68***	1

(*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$)

reported in MEGALEX (Ferrand et al., 2018), we found that the correlations between lexical variables were similar in size and significance levels, which further validates our data set. For example, and among the most widely used, word frequencies computed from books are highly correlated with other word frequencies computed from subtitles ($r = 0.78$), blogposts ($r = 0.73$), Twitter ($r = 0.63$), and newspapers ($r = 0.68$, see Table 4).

We computed correlations between our concreteness variable and those collected by Bonin et al. (2018) in French and Brysbaert, Warriner, and Kuperman (2014) and Coltheart (1981) in English. Table 5 shows that the correlations are strong and highly significant, thus ensuring the validity of the concreteness variable we collected.

Finally, to investigate the concreteness variable further, we implemented the package Ckmeans.1d.dp in R studio, an unsupervised learning algorithm for clustering univariate data (Wang & Song, 2011). Based on a Bayesian information criterion, the algorithm suggested the concreteness variable be split according to three clusters of abstractness, with cluster 1 the most abstract and cluster 3 the least abstract. The cluster variable is particularly important in relation to the previously discussed need to control the concreteness variable when manipulating semantic similarity. It is a variable which will

therefore allow experimenters to select stimuli with matching concreteness levels. We have provided the cluster variable in the [supplementary material](#).

Availability of the database

The data set for the present study is available in Excel format on the BRM and OSF websites (<https://osf.io/qsd4v/>). The main database is organised according to the following variables: word pairs in French, word-pair translation in English, word-pair mean concreteness, cluster variable based on word-pair mean concreteness, verbal association strength based on the SWOW, and mean pair similarity with associated general statistics (SD, min, max, median, range, skewness, Q1, Q3). The rest of the database is divided according to prime word and target word for the following variables: mean concreteness and associated general statistics, lexical variables (grammatical category, number of letters, and orthographic neighbours), reaction times (based on FLP and MEGALEX), and frequencies per million (movie subtitles, books, blogs, Twitter, and newspapers). The secondary database is organised following the same variables, but contains only the 70 word pairs that are semantically similar as well as verbally associated.

Table 5. Correlations of the present concreteness variable measures with those provided by other databases

Language	Study	Correlation with the present study	
		r	N
French	Bonin et al. (2018)	0.95	257
English	Coltheart (1981)	0.91	1256
English	Brysbaert et al. (2014)	0.91	1256

Note. N: number of items in common. All correlations are significant at the $p < 0.001$ level

Discussion

The present study aimed to produce French norms of semantic similarity for abstract concepts. Based on our statistical analyses, we can provide material with varying levels of semantic similarity. In addition, based on our collection of concreteness ratings and the implementation of the k-means clustering algorithm, we organised the semantic pairs according to three clusters of abstractness. Our ultimate aim is for this database to be used to design material for studies such as semantic priming studies and other language-based paradigms (see,

for example, Hutchison et al., 2013). The cross-references we computed with previously mentioned lexical databases allow stimuli to be matched on the basis of frequencies and other lexical variables. The analysis based on this cross-referencing also provides information about the potentially confounding variables that could create noise in an experimental design.

The comparison of prime and target words across the concreteness and lexical variables produced highly significant correlations, thus ensuring a good match within each pair. Further comparisons between semantic similarity and lexical variables, however, resulted in either very weak or non-significant correlations. This suggests there is no need to be particularly careful to avoid confounding lexical variables when using the similarity ratings. The strong and significant correlation in concreteness levels within word pairs, along with the cluster variable we introduced, were aimed at addressing Crutch and Jackson's (2011) suggestion that discrepancies found when studying the organisation of semantic memory according to similarity or association might be due to a binary, rather than graded, definition of concreteness levels. Indeed, when considering the organisation of semantic memory at the extremes of concrete versus abstract concepts, we lose substantial evidence for the concepts in-between these two extremes. This limitation can be addressed by considering graded levels of concreteness. Previous findings have shown that concepts are organised according to semantic similarity when concreteness increases and according to verbal association when abstractness increases.

Finally, we suggest that, when creating materials, researchers pay attention to the moderate but significant correlation between semantic similarity and the concreteness variable, insofar as results have shown that more abstract pairs are also perceived as more similar than concrete pairs.

The aim of this database was also to fill a gap in the French literature regarding norms for abstract concepts. We therefore consider the present work to be a good starting point for developing other French-language databases focusing on abstract concepts such as verbal association.

Indeed, studies using word stimuli have a tendency to focus primarily on pairing stimuli according to word frequency, word length, and age of acquisition. However, such variables fail to capture fully the effect of word processing by the human mind, as best illustrated by the percentage of variance explained in norming studies and megastudies, which stagnates between .20 and .50 (Balota, Yap, Hutchison, Cortese, Kessler, Loftis, & Treiman, 2007; Keuleers, Brysbaert, & New, 2010; Ferrand et al., 2010; Brysbaert, Mandera, & Keuleers, 2018). Newly developed variables have therefore been introduced with a view to capturing more of the word-processing phenomena. For instance, Brysbaert, Mandera, McCormick, and Keuleers (2019) introduced the word prevalence variable (the proportion of people who know a particular word), first in Dutch (Brysbaert, Stevens, Mandera, and

Keuleers, 2016; Keuleers, Stevens, Mandera, & Brysbaert, 2015), and then in English (Brysbaert et al., 2019). This variable was shown to explain an additional 6–10% of the variance in response latencies in a lexical decision task.

In addition, we consider that most norming studies have focused mainly on concrete concepts, although, as shown by Recchia and Jones (2012), abstract concepts have a richness of their own which warrants further study. For instance, Chedid, Brambati, Bedetti, Rey, Wilson, and Vallet (2019) recently introduced a perceptual strength variable for Canadian French, which aims to identify auditory and visual involvement in conceptual knowledge. In addition, the sensory experience ratings variable (SER, Juhasz, & Yap, 2013; Bonin et al., 2015, 2018) was introduced as a measure of the extent to which a word can elicit sensory and perceptual experiences. The correlation analyses between our concreteness variable and the SER variable based on the 257 items in common is 0.33. This rather low correlation goes to show that the SER variable cannot capture the same psycholinguistic phenomena as the concreteness variable, thus ensuring the relevance of the latter. We also computed the correlation between our concreteness variable and the perceptual strength variable (Chedid, Brambati, Bedetti, Rey, Wilson, & Vallet, 2019) and found that $r = 0.80$ based on 507 items in common. Although this correlation value appears rather high, it is consistent with the findings of Chedid and colleagues who reported a correlation value of $r = 0.76$ between perceptual strength and Bonin and colleagues' concreteness variables. According to Chedid et al. (2019), however, this new variable cannot be regarded as another form of concreteness since it made an independent contribution to the prediction of word latencies in word processing.

Until recently, grounding has mainly been studied in concrete concepts, owing to a previous consensus that abstract concepts are not grounded. However, several studies have shown that abstract concepts can be grounded in perceptual situations and events. In addition, Connell, Lynott, and Banks (2018) consider interoception a forgotten modality for abstract concepts and report a facilitation effect of interoceptive strength. Future work will therefore focus on developing norms that capture these modalities for abstract concepts to further our knowledge about their representation.

Conclusion

The present study aimed to provide French semantic similarity norms for 630 word pairs with varying levels of similarity and associated concreteness. The database is organised in such a way that semantic similarity and concreteness may be used as either continuous or categorical variables. The continuous variables correspond to the ratings we collected, whereas the categorical variables correspond to the cluster variable we

computed for concreteness and the median for semantic similarity. The database also provides frequency and lexical variables for matching pairs in stimuli set design. We anticipate that it will be very useful for researchers working on memory and language, especially given the growing interest for studying abstract concept representation.

Acknowledgements This research made in the FACTOLAB framework has been sponsored by Michelin Tyres Manufacturer, by the French government research programs “Investissements d’Avenir” through the IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25) and the IMobS3 Laboratory of Excellence (ANR-10-LABX-16-01).

Open practices statement In line with an open data policy, all data discussed in this article are freely available on our website on the Open Science Framework website <https://osf.io/qsd4v/>.

References

- Andrews, M., Frank, S., & Vigliocco, G. (2014). Reconciling embodied and distributional accounts of meaning in language. *Topics in Cognitive Science*, 6, 359–370. <https://doi.org/10.1111/tops.12096>
- Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, 116, 463–498. <https://doi.org/10.1037/a0016261>
- Andrews, S., Lo, S., & Xia, V. (2017). Individual differences in automatic semantic priming. *Journal of Experimental Psychology: Human Perception and Performance*, 43, 1025–1039. <https://doi.org/10.1037/xhp0000372>
- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., ... Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*. Springer New York LLC. <https://doi.org/10.3758/BF03193014>
- Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, 22, 577–660. <https://doi.org/10.1017/S0140525X99532147>
- Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358, 1177–1187. <https://doi.org/10.1098/rstb.2003.1319>
- Barsalou, L. W., Santos, A., Simmons, W.K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols embodiment, and meaning* (pp. 245–283). Oxford: University Press. <https://doi.org/10.1093/acprof:oso/9780199217274.003.0013>
- Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. Zwaan (Eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking* (pp. 129–163). New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9780511499968.007>
- Berg, T., & Levelt, W. J. M. (1990). Speaking: From intention to articulation. *The American Journal of Psychology*, 103, 409–418. <https://doi.org/10.2307/1423219>
- Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, 33, 130–174. <https://doi.org/10.1080/02643294.2016.1147426>
- Bolognesi, M., Pilgram, R., & van den Heerik, R. (2017). Reliability in content analysis: The case of semantic feature norms classification. *Behavior Research Methods*, 49, 1984–2001. <https://doi.org/10.3758/s13428-016-0838-6>
- Bonin, P., Méot, A., & Bugaïska, A. (2018). Concreteness norms for 1,659 French words: Relationships with other psycholinguistic variables and word recognition times. *Behavior Research Methods*, 50, 2366–2387. <https://doi.org/10.3758/s13428-018-1014-y>
- Bonin, P., Méot, A., Ferrand, L., & Bugaïska, A. (2015). Sensory experience ratings (SERs) for 1,659 French words: Relationships with other psycholinguistic variables and visual word recognition. *Behavior Research Methods*, 47, 813–825. <https://doi.org/10.3758/s13428-014-0503-x>
- Borghi, A. M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2019). Words as social tools: Language, sociality and inner grounding in abstract concepts. *Physics of Life Reviews*, 29, 120–153. <https://doi.org/10.1016/j.plev.2018.12.001>
- Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, 143, 263–292. <https://doi.org/10.1037/bul0000089>
- Borghi, A. M., & Pecher, D. (2011). Introduction to the special topic embodied and grounded cognition. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00187>
- Borghi, A. M., Scorolli, C., Caligiore, D., Baldassarre, G., & Tummolini, L. (2013). The embodied mind extended: Using words as social tools. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00214>
- Bruni, E., Tran, N. K., & Baroni, M. (2014). Multimodal distributional semantics. *Journal of Artificial Intelligence Research*, 49, 1–47. <https://doi.org/10.1613/jair.4135>
- Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, 27, 45–50. <https://doi.org/10.1177/0963721417727521>
- Brysbaert, M., Mandera, P., McCormick, S. F., & Keuleers, E. (2019). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods*, 51, 467–479. <https://doi.org/10.3758/s13428-018-1077-9>
- Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant’s age. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01116>
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46, 904–911. <https://doi.org/10.3758/s13428-013-0403-5>
- Buchanan, E. M., Holmes, J. L., Teasley, M. L., & Hutchison, K. A. (2013). English semantic word-pair norms and a searchable Web portal for experimental stimulus creation. *Behavior Research Methods*, 45, 746–757. <https://doi.org/10.3758/s13428-012-0284-z>
- Buchanan, E. M., Valentine, K. D., & Maxwell, N. P. (2019). English semantic feature production norms: An extended database of 4436 concepts. *Behavior Research Methods*, 51, 1849–1863. <https://doi.org/10.3758/s13428-019-01243-z>
- Carota, F., Kriegeskorte, N., Nili, H., & Pulvermüller, F. (2017). Representational similarity mapping of distributional semantics in left inferior frontal, middle temporal, and motor cortex. *Cerebral Cortex*, 27, 294–309. <https://doi.org/10.1093/cercor/bhw379>
- Chedid, G., Brambati, S. M., Bedetti, C., Rey, A. E., Wilson, M. A., & Vallet, G. T. (2019). Visual and auditory perceptual strength norms for 3,596 French nouns and their relationship with other psycholinguistic variables. *Behavior Research Methods*, 51, 2094–2105. <https://doi.org/10.3758/s13428-019-01254-w>

- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407–428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, 33, 497–505. <https://doi.org/10.1080/14640748108400805>
- Connell, L., Lynott, D., & Banks, B. (2018). Interoception: The forgotten modality in perceptual grounding of abstract and concrete concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373. <https://doi.org/10.1098/rstb.2017.0143>
- Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, 23, 371–414. https://doi.org/10.1207/s15516709cog2303_4
- Crutch, S. J. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, 128, 615–627. <https://doi.org/10.1093/brain/awh349>
- Crutch, S. J., Connell, S., & Warrington, E. K. (2009). The different representational frameworks underpinning abstract and concrete knowledge: evidence from odd-one-out judgements. *Quarterly Journal of Experimental Psychology*, 62, 1377–1388, 1388–1390. <https://doi.org/10.1080/17470210802483834>
- Crutch, S. J., & Jackson, E. C. (2011). Contrasting graded effects of semantic similarity and association across the concreteness spectrum. *Quarterly Journal of Experimental Psychology*, 64, 1388–1408. <https://doi.org/10.1080/17470218.2010.543285>
- Crutch, S. J., & Warrington, E. K. (2010). The differential dependence of abstract and concrete words upon associative and similarity-based information: Complementary semantic interference and facilitation effects. *Cognitive Neuropsychology*, 27, 46–71. <https://doi.org/10.1080/02643294.2010.491359>
- De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The “Small World of Words” English word association norms for over 12,000 cue words. *Behavior Research Methods*, 51, 987–1006. <https://doi.org/10.3758/s13428-018-1115-7>
- De Deyne, S., & Storms, G. (2008). Word associations: Norms for 1,424 Dutch words in a continuous task. *Behavior Research Methods*, 40, 198–205. <https://doi.org/10.3758/BRM.40.1.198>
- 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, 1030–1048. <https://doi.org/10.3758/BRM.40.4.1030>
- Della Rosa, P. A., Catricalà, E., Vigliocco, G., & Cappa, S. F. (2010). Beyond the abstract-concrete dichotomy: Mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words. *Behavior Research Methods*, 42, 1042–1048. <https://doi.org/10.3758/BRM.42.4.1042>
- Devereux, B. J., Tyler, L. K., Geertzen, J., & Randall, B. (2014). The Centre for Speech, Language and the Brain (CSLB) concept property norms. *Behavior Research Methods*, 46, 1119–1127. <https://doi.org/10.3758/s13428-013-0420-4>
- Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, 110, 412–431. <https://doi.org/10.1016/j.cognition.2008.11.016>
- Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology*, 1. <https://doi.org/10.3389/fpsyg.2010.00242>
- Dove, G. (2014). Thinking in words: Language as an embodied medium of thought. *Topics in Cognitive Science*, 6, 371–389. <https://doi.org/10.1111/tops.12102>
- Duñabeitia, J. A., Avilés, A., Afonso, O., Scheepers, C., & Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition*, 110, 284–292. <https://doi.org/10.1016/j.cognition.2008.11.012>
- Ferrand, L. (2001). Normes d’associations verbales pour 260 mots « abstraits » [Word association norms for 260 “abstract” words]. *L’Année Psychologique*, 101, 683–721. <https://doi.org/10.3406/psy.2001.29575>
- Ferrand, L., & Alario, F. X. (1998). Word association norms for 366 names of objects. [Word association norms for 366 concrete objects words]. *L’Année Psychologique*, 98, 659–709. <https://doi.org/10.3406/psy.1998.28564>
- Ferrand, L., Méot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., ... Grainger, J. (2018). MEGALEX: A megastudy of visual and auditory word recognition. *Behavior Research Methods*, 50, 1285–1307. <https://doi.org/10.3758/s13428-017-0943-1>
- Ferrand, L., & New, B. (2003). Associative and semantic priming in the mental lexicon. In P. Bonin (Ed.), *The mental lexicon: Some words to talk about words* (pp. 25–43). New York: Nova Science Publishers.
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., ... Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudo words. *Behavior Research Methods*, 42, 488–496. <https://doi.org/10.3758/BRM.42.2.488>
- Ferré, P., Guasch, M., García-Chico, T., & Sánchez-Casas, R. (2015). Are there qualitative differences in the representation of abstract and concrete words? Within-language and cross-language evidence from the semantic priming paradigm. *Quarterly Journal of Experimental Psychology*, 68, 2402–2418. <https://doi.org/10.1080/17470218.2015.1016980>
- Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, 44, 516–547. <https://doi.org/10.1006/jmla.2000.2728>
- Firth, J. R. (1957). Applications of general linguistics. *Transactions of the Philological Society*, 56, 1–14. <https://doi.org/10.1111/j.1467-968X.1957.tb00568.x>
- Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition*, 8, 263–367. [https://doi.org/10.1016/0010-0277\(80\)90008-6](https://doi.org/10.1016/0010-0277(80)90008-6)
- Gallese, V., & Lakoff, G. (2005). The brain’s concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22, 455–479. <https://doi.org/10.1080/02643290442000310>
- Geng, J., & Schnur, T. T. (2015). The representation of concrete and abstract concepts: Categorical versus associative relationships. *Journal of Experimental Psychology: Learning Memory and Cognition*, 41, 22–41. <https://doi.org/10.1037/a0037430>
- Gimenes, M., & New, B. (2016). Worldlex: Twitter and blog word frequencies for 66 languages. *Behavior Research Methods*, 48, 963–972. <https://doi.org/10.3758/s13428-015-0621-0>
- Glenberg, A. M. (1997). Mental models, space, and embodied cognition. In T. B. Ward & S. M. Smith (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 495–522). Washington, DC: American Psychological Association.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, 9, 558–565. <https://doi.org/10.3758/BF03196313>
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, 114, 211–244. <https://doi.org/10.1037/0033-295X.114.2.211>
- Hamilton, A. C., & Coslett, H. B. (2008). Refractory access disorders and the organization of concrete and abstract semantics: Do they differ? *Neurocase*, 14, 131–140. <https://doi.org/10.1080/13554790802032218>
- Harpaintner, M., Trumpp, N. M., & Kiefer, M. (2018). The semantic content of abstract concepts: A property listing study of 296 abstract

- words. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.01748>
- Harris, Z. S. (1954). Distributional Structure. *Distributional Structure, Word*, 10, 146–162. <https://doi.org/10.1080/00437956.1954.11659520>
- Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control, and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review*, 125, 293–328. <https://doi.org/10.1037/rev0000094>
- Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin and Review*, 10, 785–813. <https://doi.org/10.3758/BF03196544>
- Hutchison, K. A., Balota, D. A., Cortese, M. J., & Watson, J. M. (2008). Predicting semantic priming at the item level. *Quarterly Journal of Experimental Psychology*, 61, 1036–1066. <https://doi.org/10.1080/17470210701438111>
- Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E. R., Tse, C. S., ... Buchanan, E. (2013). The semantic priming project. *Behavior Research Methods*, 45, 1099–1114. <https://doi.org/10.3758/s13428-012-0304-z>
- Juhász, B. J., & Yap, M. J. (2013). Sensory experience ratings for over 5,000 mono- and disyllabic words. *Behavior Research Methods*, 45, 160–168. <https://doi.org/10.3758/s13428-012-0242-9>
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42, 643–650. <https://doi.org/10.3758/BRM.42.3.643>
- Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology*, 68, 1665–1692. <https://doi.org/10.1080/17470218.2015.1022560>
- Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, 48, 805–825. <https://doi.org/10.1016/j.cortex.2011.04.006>
- Kim, S. Y., Yap, M. J., & Goh, W. D. (2019). The role of semantic transparency in visual word recognition of compound words: A megastudy approach. *Behavior Research Methods*, 51, 2722–2732. <https://doi.org/10.3758/s13428-018-1143-3>
- Kintsch, W., McNamara, D. S., Dennis, S., Landauer, T. K., McNamara, D. S., Dennis, S., & Landauer, T. K. (2007). LSA and meaning: in theory and application, 479–492. <https://doi.org/10.4324/9780203936399-32>
- Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140, 14–34. <https://doi.org/10.1037/a0021446>
- Kremer, G., & Baroni, M. (2011). A set of semantic norms for German and Italian. *Behavior Research Methods*, 43, 97–109. <https://doi.org/10.3758/s13428-010-0028-x>
- Lakoff, G., & Johnson, M. (1980). Conceptual metaphor in everyday language. *The Journal of Philosophy*, 77, 453. <https://doi.org/10.2307/2025464>
- 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, 211–240. <https://doi.org/10.1037/0033-295X.104.2.211>
- Lebani, G. E., Bondielli, A., & Lenci, A. (2015). You are what you do: An empirical characterization of the semantic content of the thematic roles for a group of Italian verbs. *Journal of Cognitive Science*, 16, 399–428.
- Lenci, A. (2008). Distributional semantics in linguistic and cognitive research. *Italian Journal of Linguistics*, 20, 1–31.
- Lenci, A. (2018). Distributional models of word meaning. *Annual Review of Linguistics*, 4, 151–171. <https://doi.org/10.1146/annurev-linguistics-030514-125254>
- Lenci, A., Baroni, M., Cazzolli, G., & Marotta, G. (2013). BLIND: A set of semantic feature norms from the congenitally blind. *Behavior Research Methods*, 45, 1218–1233. <https://doi.org/10.3758/s13428-013-0323-4>
- Lenci, A., Lebani, G. E., & Passaro, L. C. (2018). The emotions of abstract words: A distributional semantic analysis. *Topics in Cognitive Science*, 10, 550–572. <https://doi.org/10.1111/tops.12335>
- Levelt, W. J. M., Roelofs, A., & Meyer, S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38. <https://doi.org/10.1017/s0140525x99001776>
- Louwerse, M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In *Symbols and Embodiment Debates on meaning and cognition* (pp. 309–326). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199217274.003.0015>
- Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin and Review*, 15, 838–844. <https://doi.org/10.3758/PBR.15.4.838>
- Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, 3, 273–302. <https://doi.org/10.1111/j.1756-8765.2010.01106.x>
- Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, 114, 96–104. <https://doi.org/10.1016/j.cognition.2009.09.002>
- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, 28, 203–208.
- Machery, E. (2016). The amodal brain and the offloading hypothesis. *Psychonomic Bulletin and Review*, 23, 1090–1095. <https://doi.org/10.3758/s13423-015-0878-4>
- 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. <https://doi.org/10.1016/j.jphysparis.2008.03.004>
- Maki, W. S., Krinsky, M., & Muñoz, S. (2006). An efficient method for estimating semantic similarity based on feature overlap: Reliability and validity of semantic feature ratings. *Behavior Research Methods*, 38, 153–157. <https://doi.org/10.3758/BF03192761>
- Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, 92, 57–78. <https://doi.org/10.1016/j.jml.2016.04.001>
- Mate, J., Allen, R. J., & Baqués, J. (2012). What you say matters: exploring visual-verbal interactions in visual working memory. *Quarterly Journal of Experimental Psychology*, 65, 395–400. <https://doi.org/10.1080/17470218.2011.644798>
- McNamara, T. P. (1992). Theories of Priming: I. Associative Distance and Lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1173–1190. <https://doi.org/10.1037/0278-7393.18.6.1173>
- McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning Memory and Cognition*, 24, 558–572. <https://doi.org/10.1037/0278-7393.24.3.558>
- 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, 547–559. <https://doi.org/10.3758/BF03192726>
- McRae, K., De Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126, 99–130. <https://doi.org/10.1037/0096-3445.126.2.99>

- Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*, 788–804. <https://doi.org/10.1016/j.cortex.2010.11.002>
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234. <https://doi.org/10.1037/h0031564>
- Moldovan, C. D., Ferré, P., Demestre, J., & Sánchez-Casas, R. (2015). Semantic similarity: normative ratings for 185 Spanish noun triplets. *Behavior Research Methods*, *47*, 788–799. <https://doi.org/10.3758/s13428-014-0501-z>
- New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*, 661–677. <https://doi.org/10.1017/S014271640707035X>
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, and Computers*, *36*, 516–524. <https://doi.org/10.3758/BF03195598>
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). A lexical database for contemporary French on Internet : LEXIQUE. *L'Année Psychologique*, *101*, 447–462. <https://doi.org/10.3406/psy.2001.1341>
- Nishiyama, R. (2013). Dissociative contributions of semantic and lexical-phonological information to immediate recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 642–648. <https://doi.org/10.1037/a0029160>
- Oliveira, J., Perea, M. V., Ladera, V., & Gamito, P. (2013). The roles of word concreteness and cognitive load on interhemispheric processes of recognition. *Laterality*, *18*, 203–215. <https://doi.org/10.1080/1357650X.2011.649758>
- Ostarek, M., & Huettig, F. (2019). Six challenges for embodiment research. *Current Directions in Psychological Science*, *28*, 593–599. <https://doi.org/10.1177/0963721419866441>
- Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*, 1–25. <https://doi.org/10.1037/h0025327>
- Pecher, D. (2018). Curb your embodiment. *Topics in Cognitive Science*, *10*, 501–517. <https://doi.org/10.1111/tops.12311>
- Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research*, *66*, 180–194. <https://doi.org/10.1007/s00426-002-0086-5>
- Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. Cognitive Science Society (Ed.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 37–42). Hillsdale, NJ: Erlbaum
- Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review*, *107*, 786–823. <https://doi.org/10.1037/0033-295X.107.4.786>
- Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, *17*, 458–470. <https://doi.org/10.1016/j.tics.2013.06.004>
- Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience*, *17*, 884–892. <https://doi.org/10.1162/0898929054021111>
- Qualtrics (2020) Qualtrics.com. Available at: <http://www.qualtrics.com/>
- Recchia, G., & Jones, M. N. (2012). The semantic richness of abstract concepts. *Frontiers in Human Neuroscience*, *6*. <https://doi.org/10.3389/fnhum.2012.00315>
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284. [https://doi.org/10.1016/S0010-0277\(97\)00027-9](https://doi.org/10.1016/S0010-0277(97)00027-9)
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Sánchez-Casas, R., Ferré, P., García-Albea, J. E., & Guasch, M. (2006). The nature of semantic priming: Effects of the degree of semantic similarity between primes and targets in Spanish. *European Journal of Cognitive Psychology*, *18*, 161–184. <https://doi.org/10.1080/09541440500183830>
- Schwanenflugel, P. J., Hamishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, *27*, 499–520. [https://doi.org/10.1016/0749-596X\(88\)90022-8](https://doi.org/10.1016/0749-596X(88)90022-8)
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*, 214–241. <https://doi.org/10.1037/h0036351>
- Sperber, A. D., Devellis, R. F., & Boehlecke, B. (1994). Cross-cultural translation. *Journal of Cross-Cultural Psychology*, *25*, 501–524. <https://doi.org/10.1177/0022022194254006>
- Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, *1*, 219–247. <https://doi.org/10.1515/langcog.2009.011>
- Vigliocco, G., & Vinson, D. P. (2007). Semantic representation. In G. Gaskell & G. Altmann (Eds.), *The Oxford handbook of psycholinguistics* (pp. 195–215). Oxford, England: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198568971.013.0012>
- 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. <https://doi.org/10.1016/j.cogpsych.2003.09.001>
- Villani, C., Lugli, L., Liuzza, M. T., & Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition*, *11*, 403–430. <https://doi.org/10.1017/langcog.2019.23>
- Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*, 183–190. <https://doi.org/10.3758/BRM.40.1.183>
- Vivas, J., Vivas, L., Comesaña, A., Coni, A. G., & Vorano, A. (2017). Spanish semantic feature production norms for 400 concrete concepts. *Behavior Research Methods*, *49*, 1095–1106. <https://doi.org/10.3758/s13428-016-0777-2>
- Wang, H. & Song, M. (2011). Ckmeans.1d.dp: optimal k-means clustering in one dimension by dynamic programming. *R Journal*, *3*, 29–33.
- Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science*, *29*, 719–736. https://doi.org/10.1207/s15516709cog0000_33
- Wu, L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, *132*, 173–189. <https://doi.org/10.1016/j.actpsy.2009.02.002>
- Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin and Review*, *15*, 971–979. <https://doi.org/10.3758/PBR.15.5.971>
- Zwaan, R. A. (2004). The immersed experienter: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation - Advances in Research and Theory*, *44*, 35–62. [https://doi.org/10.1016/S0079-7421\(03\)44002-4](https://doi.org/10.1016/S0079-7421(03)44002-4)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.