



# Levels of Abstractness in Semantic Noun and Verb Processing: The Role of Sensory-Perceptual and Sensory-Motor Information

Jet M. J. Vonk<sup>1</sup> · Loraine K. Obler<sup>2</sup> · Roel Jonkers<sup>3</sup>

Published online: 2 January 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

Effects of concreteness and grammatical class on lexical-semantic processing are well-documented, but the role of sensory-perceptual and sensory-motor features of concepts in underlying mechanisms producing these effects is relatively unknown. We hypothesized that processing dissimilarities in accuracy and response time performance in nouns versus verbs, concrete versus abstract words, and their interaction can be explained by differences in semantic weight—the combined amount of sensory-perceptual and sensory-motor information to conceptual representations—across those grammatical and semantic categories. We assessed performance on concrete and abstract subcategories of nouns and verbs with a semantic similarity judgment task. Results showed that when main effects of concreteness and grammatical class were analyzed in more detail, the grammatical-class effect, in which nouns are processed more accurately and quicker than verbs, was only present for concrete words, not for their abstract counterparts. Moreover, the concreteness effect, measured at different levels of abstract words, was present for both nouns and verbs, but it was less pronounced for verbs. The results do not support the grammatical-class hypothesis, in which nouns and verbs are separately organized, and instead provide evidence in favor of a unitary semantic space, in which lexical-semantic processing is influenced by the beneficial effect of sensory-perceptual and sensory-motor information of concepts.

**Keywords** Grammatical class · Imageability · Concreteness · Semantics · Comprehension · Mental lexicon

---

✉ Jet M. J. Vonk  
jv2528@cumc.columbia.edu

<sup>1</sup> Department of Neurology, Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University Medical Center, 622 W 168th St, New York, NY 10032-3784, USA

<sup>2</sup> Department of Speech-Language-Hearing Sciences, The Graduate Center of the City University of New York, New York, NY, USA

<sup>3</sup> Department of Linguistics, University of Groningen, Groningen, The Netherlands

## Introduction

The concepts underlying the words in our lexicon are composed of semantically meaningful features, including various sensory, motor, linguistic, and affective features. Common features and properties, especially sensory-perceptual and sensory-motor ones, define relations between concepts, such as membership of a semantic category. Across categories, as well as within categories, conceptual representations vary in how salient a sensory-perceptual or sensory-motor feature is to its meaning (i.e., its semantic weight, Vinson et al. 2003). For example, a highly salient and defining feature of *animals* is *biological motion*, while *manipulation* and *mechanical motion* are salient features of *tools* (Martin 2007). Sensory-perceptual and sensory-motor features therefore tend to only be salient for concrete words, and not for abstract ones. Concreteness reflects the degree to which the meaning of the concept that the word refers to is tangible and sensorily perceptible (Brysbaert et al. 2014). In other words, concreteness ratings are an indicator of the combined weight of sensory-perceptual and sensory-motor features (Vinson et al. 2003). Many studies have demonstrated that concrete words are easier to process than abstract words in a variety of cognitive and linguistic tasks (e.g., Adorni and Proverbio 2012; James 1975; but see Kousta et al. 2011), suggesting that the richer sensory-perceptual and sensory-motor representation of concrete words may facilitate their lexical processing (Hoffman 2016).

While the concreteness effect is one of the most robust phenomena in psycholinguistic research, it has mainly been investigated in nouns, as opposed to verbs. The few studies that included both nouns and verbs in their investigation of the concreteness effect found an interaction between grammatical class and concreteness (e.g., Lee and Federmeier 2008). The results are mixed across investigations, partly due to the breadth of techniques and outcome variables used; however, an overall theme that has emerged across these studies is that a concreteness effect is often present in verbs, yet it is less pronounced than in nouns (e.g., Kellenbach et al. 2002; Zhang et al. 2006).

In general, and independent of concreteness, differences across grammatical class, specifically between nouns and verbs, have commonly been found with either behavioral methods (e.g., Barresi et al. 2000) or neuroimaging techniques (e.g., Tranel et al. 2005; Warburton et al. 1996). Accuracy and response time (RT) differences, in which nouns are usually processed better than verbs, have often been observed in those with neurological impairment due to, e.g., stroke or dementia (e.g., Bak and Hodges 2003; Berndt et al. 2002; Jonkers 1998). Such discrepancies in performance between nouns and verbs on a variety of tasks have led some to argue that the representations of nouns and verbs are separate and functionally independent (Crepaldi et al. 2011). For example, some researchers have argued that lexical forms (i.e., grammatical class) are an organizational principle for knowledge of language in the brain (e.g., Caramazza and Hillis 1991; Hillis and Caramazza 1995; Shapiro et al. 2000; Silveri and Di Betta 1997) and that there are separate neural substrates for the two grammatical classes (e.g., Daniele et al. 1994). This theory would thus predict that the difference between nouns and verbs applies not only to concrete nouns (i.e., objects) and concrete verbs (i.e., actions), but also to their abstract counterparts.

Others argue that dissociations in performance between nouns and verbs are based on semantic features, explaining the results found so far as not based on a distinction in grammatical class, but on the distinction between the sensory-perceptual and sensory-motor features of objects (represented by concrete nouns) and actions/events (represented by concrete verbs) (e.g., Barber et al. 2010; Breedin et al. 1998; Moseley and Pulvermüller 2014; Pulvermüller et al. 1999; Siri et al. 2007; Vigliocco et al. 2004, 2006; Vinson

and Vigliocco 2002; Vonk 2015). Vinson et al. (2003) showed that objects generally have greater featural weight than actions. Additionally, it has been shown that nouns referring to actions (e.g., murder) are more similar to verbs referring to actions than to nouns referring to objects (Pulvermüller et al. 1999; Vinson and Vigliocco 2002), thereby providing evidence against the grammatical-class hypothesis and in favor of a unitary semantic space, in which concepts are organized based on semantically meaningful features, such as sensory, motor, linguistic, and affective information (Vigliocco et al. 2004). This view, thus, would predict that differences in the weight of features associated with concepts may result in differences in performance between concrete nouns and verbs (because, e.g., visual features are very salient for objects and less so for actions, Vigliocco et al. 2004), but that in the near absence of such features, performance on abstract nouns and verbs should be relatively equal.

Studies on noun versus verb processing often neglect semantic differences between objects and actions (Kemmerer 2014; Vigliocco et al. 2011). As a result, nouns and verbs are frequently defined only by their semantic membership (i.e., object or action), as opposed to also including their other elements that underlie grammatical class membership following the general principles of linguistics (Croft 2000). The current study specifically investigated joint effects of concreteness and grammatical class on semantic processing. To do so, we deliberately included nouns and verbs that were neither objects nor actions (for discussion of the implications of linguistic typology for neurolinguistics research, see Kemmerer 2014).

We investigated differences in processing nouns and verbs across concrete words and multiple levels of abstractness. As recent studies suggested that differences in processing nouns versus verbs are regulated by featural weight (Buzzeo et al. 2017; Vinson et al. 2003), we hypothesized that categories of concrete nouns (i.e., *animals*, *furniture*, and *tools*) are processed more accurately and quickly than categories of concrete verbs (i.e., *change of state*, *hitting*, and *cutting*). Moreover, we hypothesized that in the absence of sensory-perceptual and sensory-motor features, abstract nouns and verbs are processed with equal accuracy and RT, counter to the grammatical-class hypothesis and in favor of a unitary semantic space. Concreteness is often studied as a dichotomous classification between concrete and abstract, but to assess the effect of a gradual decrease of sensory-perceptual and sensory-motor features across concepts, we included multiple levels of abstractness using categories of *mildly*, *moderately*, and *highly* abstract words.

## Methodology

### Participants

We tested 16 native Dutch speakers<sup>1</sup> with a mean education of 14.2 years (SD=3.4), aged 50–68 (mean=57.8, SD=6.4) of whom 10 were males. All participants reported having normal or corrected-to-normal vision and no history of learning disabilities, neurological problems, or head injury.

---

<sup>1</sup> We treated all participants as effectively monolingual since they were originally and currently from Noord-Brabant which is a monolingual region and individuals of the generation of our participants did not have the broad exposure to other languages that educational reforms, television, and internet have provided younger generations.

## Materials

We created a semantic similarity judgment task, similar in structure to the Pyramids and Palm Trees test (Howard and Patterson 1992), to assess semantic processing of twelve different concrete and abstract noun and verb categories. Each item consisted of a visual word triad with one word centered in the top line and two words on a second line. The participant had to decide which of the two words in the second line was closer in meaning to the one on top (e.g., *horse: donkey – goat*). This task was selected to avoid task-related effects on semantic performance due to complex visual processing required for most picture tasks or lexical-phonological retrieval required for production tasks, and because it ensures deep semantic processing (Sabsevitz et al. 2005).

Table 1 provides an overview of the test stimuli including examples. A total of 204 triads were tested, of which half were noun triads and the other half verb triads. Half of the noun triads and half of the verb triads tested were concrete and half were abstract. The 51 concrete noun triads were subdivided into the three categories of 17 items: *animals*, *tools/instruments*, and *furniture*, and the concrete verb triads similarly into the three categories: *change of state*, *hitting*, and *cutting*. These concreteness categories were based on the stimuli of Sabsevitz et al. (2005) for nouns and Kemmerer et al. (2008) for verbs. For both nouns and verbs, the abstract triads were subdivided into *mildly abstract*, *moderately abstract*, and *highly abstract*, a classification used by Crutch and Warrington (2004). Each category had 17 triads.

Every triad consisted of only concrete words or only abstract words, and every triad consisted of only words from the same category, e.g., all three words in a triad were *animals*, or all three were *mildly abstract*. Only words that occurred at least twice in a million in the WebCelex database (Max Planck Institute for Psycholinguistics 2001) were included and no words shorter than four letters were included (max. 11 letters). Due to the restrictions in semantically valid options for triads, other psycholinguistic variables could not be controlled for in the design of the stimuli; yet, for each triad the mean lexical frequency (Keuleers et al. 2010) was calculated to covary for in the statistical analyses. Table 2 reflects the mean log frequency of the test stimuli per category.

Concreteness ratings were derived from the norms for the Dutch language by Van Loon-Vervoorn (1985), incorporated in the online WebCelex database. Each word within a triad had a concreteness rating that fell within the concreteness range of the category it belonged to. All concrete words had a concreteness rating of 5 or higher on a 7-point scale. The *mildly abstract* words had a rating between 4 and 5 on this 7-point scale (e.g., *gebied: buurt—detail*; translated into English: “area: neighborhood—detail”), the *moderately abstract* words a rating between 3 and 4 (e.g., *aanbod: keus—beheer*; translated into English: “offer: option—management”), and the *highly abstract* words a rating between 1 and 3 (e.g., *macht: bewind—cultuur*; translated into English: “power: control—culture”). The concreteness rating for each triad was calculated as the average of the ratings for the three words within it (Table 2).

During development of the materials, the triads of words were scored for difficulty of judging semantic relatedness in a pilot rating study. For each semantic category, 30 triads were developed, resulting in a total of 180 triads. In the pilot study, 22 native-Dutch-speaking participants (19 female, mean age = 33.7, SD = 17.5, range = 22–65 years; no overlap in participants with the experimental task) made a difficulty judgment on a set of triads, such that every triad was rated 17 times in total. The participants were asked to rate the difficulty of judging the correct answer for each triad on a Likert scale from 1 (extremely easy) to 7

**Table 1** Overview of test stimuli

Total	Grammatical category	Concreteness	Semantic category	Example (given word: target – distracter) + [translation]
204 triads	Nouns (102)	Concrete (51)	Animals (17)	panter: tijger – buffel [panther: tiger – buffalo]
			Furniture (17)	divan: bank – kruk [settee: couch – stool]
		Abstract (51)	Tools/instruments (17)	naald: speld – zwaard [needle: pin – sword]
			Mildly abstract (17)	gemeente: evenement – bevolking [community: event – population]
			Moderately abstract (17)	effect: gevolg – bijstand [effect: result – assistance]
			Highly abstract (17)	gedachte: idee – kwestie [thought: idea – issue]
	Verbs (102)	Concrete (51)	Change of state (17)	breken: knappen – zwellen [to break: to snap – to swell up]
			Hitting (17)	kloppen: bonken – duwen [to knock: to bang – to push]
			Cutting (17)	pellen: schillen – prikken [to peel: to skin – to prick]
		Abstract (51)]	Mildly abstract (17)	luchten: tochten – kwetsen [to vent: to let the wind through – to offend]
			Moderately abstract (17)	kennen: weten – eisen [to know: to be informed – to demand]
			Highly abstract (17)	menen: bedoelen – dulden [to be of the opinion: to mean – to allow]

(n) number of triads in that category

**Table 2** Mean (SD) psycholinguistic properties and accuracy/response time per category

Category	Concreteness	Log frequency	Accuracy	Response time
<i>Nouns</i>				
Animals	6.28 (0.29)	2.24 (0.43)	96.5 (18.5)	3043 (1973)
Furniture	6.39 (0.15)	2.21 (0.54)	98.5 (12.1)	2420 (1426)
Tools/instruments	6.44 (0.26)	2.17 (0.31)	98.5 (12.1)	2636 (1604)
Mildly abstract	4.51 (0.17)	2.71 (0.40)	97.4 (16.0)	3262 (2061)
Moderately abstract	3.46 (0.13)	2.49 (0.27)	95.1 (21.6)	3499 (1999)
Highly abstract	2.62 (0.19)	2.54 (0.56)	91.2 (28.4)	3801 (2438)
<i>Verbs</i>				
Change of state	5.88 (0.34)	2.23 (0.29)	94.0 (23.8)	3434 (2098)
Hitting	6.19 (0.28)	2.41 (0.27)	93.3 (25.0)	3342 (1957)
Cutting	6.16 (0.19)	2.08 (0.47)	92.3 (26.3)	3055 (1700)
Mildly abstract	4.54 (0.13)	2.59 (0.62)	97.0 (17.0)	3151 (1894)
Moderately abstract	3.49 (0.20)	3.08 (0.74)	96.6 (18.1)	3408 (2080)
Highly abstract	2.50 (0.19)	2.25 (0.59)	94.7 (22.4)	3373 (2138)
			89.6 (30.6)	4300 (2529)

Accuracy in percentage, response time in milliseconds; *SD* standard deviation; concreteness and frequency values for a test item are calculated as the average of the values for the three words within that triad; concreteness values on a 1–7 scale were derived from the norms by Van Loon-Vervoorn (1985)

(impossible). For every triad, the mean and median ratings were calculated; the criterion for selection was a score lower than or equal to 4 (average difficulty) on both measures of central tendency. The average difficulty of selected triads was close to equal across categories (mean total triad difficulty = 2.67, *SD* = .20, range = 2.30–3.09; median total triad difficulty = 2.38, *SD* = .29, range = 1.85–2.88).

## Procedure

The materials were pseudo-randomly organized as follows. Within each semantic category, there was an even distribution between the correct answer appearing on the right or left side (9 vs. 8 out of 17). In the experiment, no correct match to the top word appeared on one side (left or right) more than three times in a row. No semantic category appeared more than two times in a row. Certain words (22% of the total words used) appeared more than one time in the materials, but always in a different triad-combination and the same word did not occur in a triad that was less than three triads distant. The triads were divided into blocks of 10 each, apart from two blocks containing 11 triads.

The task was performed on a laptop, model HP Pavilion dv7-3030ed Entertainment Notebook PC (VL101EA#ABH), AMD Turion™ II Dual-Core Mobile M500 (2.20 GHz, 4.00 GB of RAM), 17.3" screen with the system Windows 7 Home Premium (Version 2009). The experiment was run via E-prime V2.0 (2.0.8.22) software that measured both accuracy and RT in milliseconds (Schneider et al. 2002).

The participants were tested individually in a quiet room seated at a table with a laptop in front of them. Instructions were given on the screen, and the researcher was seated next to the participant to answer any questions. Participants were instructed to decide which of the two bottom words was closest in meaning to the word on top and to use their index fingers to press

the button corresponding to their answer. If they chose the left-hand word as the semantic match to the top word, they were to press the blue button on the left (a blue sticker covered the 'z' key); if they chose the right-hand word, they pressed the yellow button on the right (a yellow sticker covered the '/' key). Participants were free to take breaks between every block but were requested not to pause or talk during a block. A short break was scheduled after the first and second thirds of the experiment. Including these breaks, the duration of the experiment was approximately 30 min.

## Statistical Analysis

We excluded 62 out of 3264 triads from analysis. An item analysis excluded triads for which a participant took longer than three standard deviations from the mean RT (i.e., > 12,713 ms; deleted  $n=61$ ). Additionally, one triad was excluded on an individual basis because of a moment of distraction during testing unrelated to the experiment. Due to the typical positively skewed distribution of RT, a natural logarithmic transformation was applied to obtain a normally distributed RT variable. Analyses of RT were performed on accurately answered triads only.

We used descriptive statistics to derive the distributional characteristics of demographic and performance variables. Generalized linear mixed models were used to analyze the relationships between grammatical and semantic categories and the relationship among levels of abstractness as measured by accuracy and RT. In models to analyze the relationship between grammatical and semantic categories, we entered grammatical class, concreteness, and their interaction as fixed effects, together with the covariates age, gender, years of education, and mean lexical frequency. Models included a random intercept for subjects to account for inter-personal variability. Model convergence was not reached when including a by-subject random slope for grammatical class and concreteness and their interaction.

In models to analyze the relationship among levels of abstractness, we restricted our sample to include only the categories of interest for the comparison in question. We entered category membership and the covariates age, gender, years of education, and mean lexical frequency as fixed effects. As random effects, the model included a random intercept for subjects. Model convergence was not reached when including a by-subject random slope for category membership. Multiple pairwise comparisons to analyze effects within category membership groups (i.e., the three concrete noun categories, the three concrete verb categories, the three abstract noun categories, and the three abstract verb categories) and hypothesized comparisons between categories across category membership groups used the sequential Šidák correction for multiple comparisons.

RT, a continuous measurement, was analyzed with linear mixed models fitting a random intercept and fixed slope. Apart from the distribution and link in comparison to the accuracy models, the RT models were built with the same fixed effects as described in the accuracy models. All analyses were performed in IBM SPSS Statistics Version 23 (IBM Corp. 2015).

## Results

### Participant Characteristics

Participants scored on average 95.2% (SD=.04) correct, ranging between 85.3% and 98.5%. The average RT per item across participants was 3252 ms (SD=838), ranging

between 1994 and 4914 ms. Table 2 displays the means and standard deviations of accuracy and RT performance on nouns versus verbs, and the various semantic subcategories within nouns and verbs. Figure 1 provides a visual overview of the participants' performance on the semantic categories.

### Grammatical Class and Concreteness Effects

A main effect of grammatical category revealed a more accurate ( $F(1, 3189)=9.971, p=.002$ ) and quicker ( $F(1, 3036)=53.867, p<.001$ ) performance on noun than verb triads. Analyzed as a dichotomous variable, a main effect of concreteness indicated a more accurate ( $F(1, 3189)=4.803, p=.028$ ) and quicker ( $F(1, 3036)=129.000, p<.001$ ) performance on concrete than abstract triads. Moreover, effects of grammatical class and concreteness interacted for both accuracy ( $F(1, 3188)=8.675, p=.003$ ) and RT performance ( $F(1, 3035)=26.828, p<.001$ ; Fig. 2). In detail, participants performed more accurate ( $t(3188)=2.804, p=.005$ ) and quicker ( $t(3035)=-11.912, p<.001$ ) on concrete than abstract nouns while this concreteness effect was not present for verbs in accuracy performance ( $t(3188)=-.006, p=.995$ ) and to a much lesser extent in RT performance ( $t(3035)=-4.698, p<.001$ ). Similarly, performance on concrete noun triads was better than on concrete verb triads in both accuracy ( $t(3188)=3.043, p=.002$ ) and RT performance ( $t(3035)=-8.892, p<.001$ ) while this grammatical effect was not present for abstract words in either accuracy ( $t(3188)=.615, p=.539$ ) or RT performance ( $t(3035)=-1.451, p=.147$ ).

### Semantic Categories and Levels of Abstractness

Accuracy and RT performance was comparable among concrete categories but declined the more abstract the triads became. We found no main effects in models restricted to either the concrete noun categories, i.e., *animals*, *furniture*, and *tools/instruments* for accuracy ( $F(2, 799)=.048, p=.953$ ) and RT performance ( $F(2, 785)=2.194, p=.112$ ) or the concrete verb categories, i.e., *change of state*, *hitting*, and *cutting* for accuracy performance ( $F(2, 794)=1.699, p=.183$ ). Concrete verb categories differed in RT performance ( $F(2, 748)=5.153, p=.006$ ), as *hitting* verb triads were responded to faster than *change of state* verb triads ( $t(748)=-3.144, p=.005$ ).

In contrast, a main effect of abstract noun categories was observed in both accuracy ( $F(2, 779)=4.435, p=.012$ ) and RT performance ( $F(2, 736)=4.820, p=.008$ ), in which participants were faster to respond to mildly abstract noun triads than to highly abstract noun triads ( $t(736)=-2.995, p=.008$ ). A similar pattern was observed for accuracy performance, although not significant because of multiple comparison correction. As well, a main effect of abstract verb categories was observed for both accuracy ( $F(2, 778)=8.729, p<.001$ ) and RT performance ( $F(2, 728)=17.077, p<.001$ ). Compared to highly abstract verb triads, performance was better on mildly abstract (accuracy:  $t(778)=2.754, p=.018$ ; RT:  $t(728)=-5.012, p<.001$ ) and moderately abstract verb triads (accuracy:  $t(778)=2.632, p=.018$ ; RT:  $t(728)=5.287, p<.001$ ). In a model with all abstract categories (i.e., both nouns and verbs), comparisons across abstract noun and verb pairs showed that performance on *mildly abstract* noun triads is not different from that on *mildly abstract* verb triads (accuracy:  $t(1567)=-.602, p=.907$ ; RT:  $t(1474)=-1.241, p=.702$ ), neither is the performance on *moderately abstract* noun versus *moderately abstract* verb triads (accuracy:  $t(1567)=.443, p=.907$ ; RT:  $t(1474)=1.898, p=.341$ ), nor is the performance



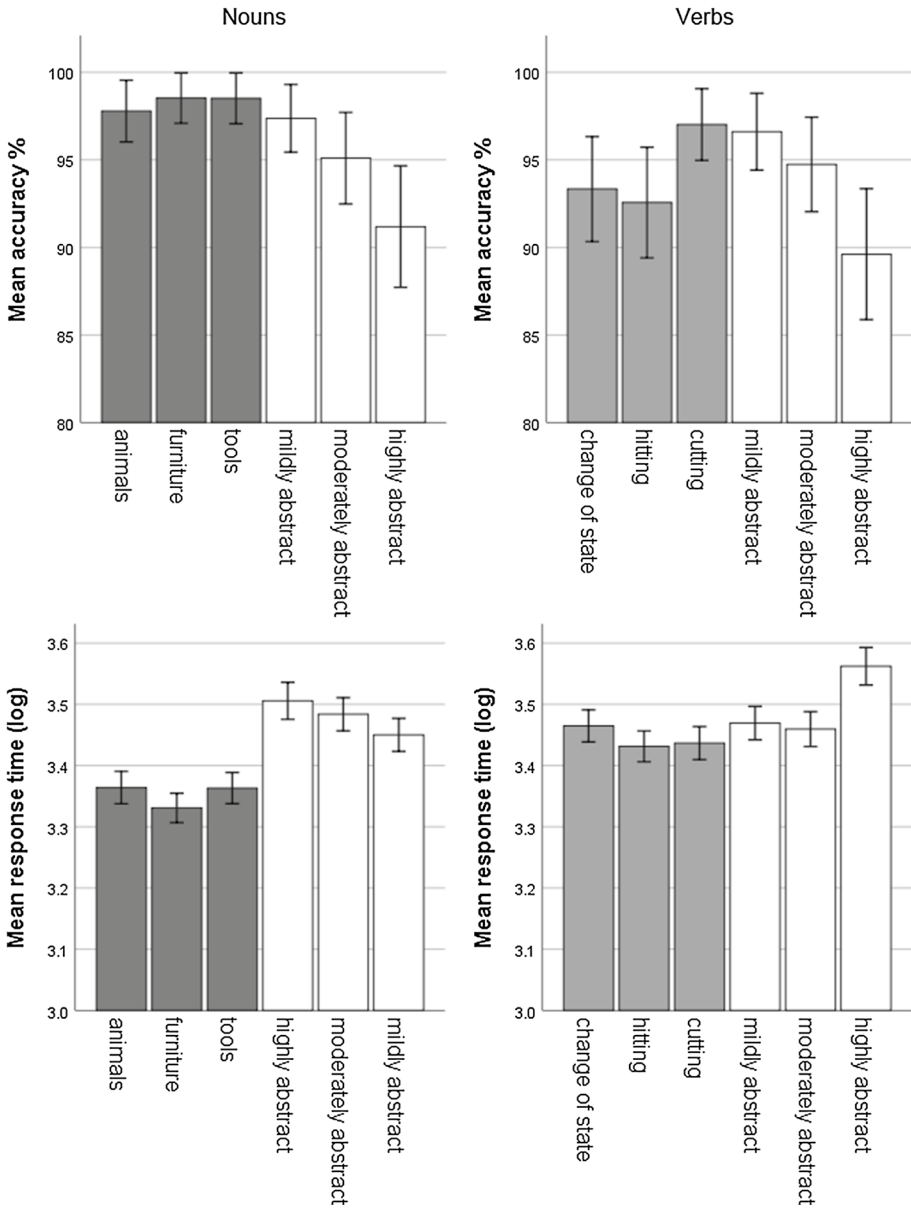
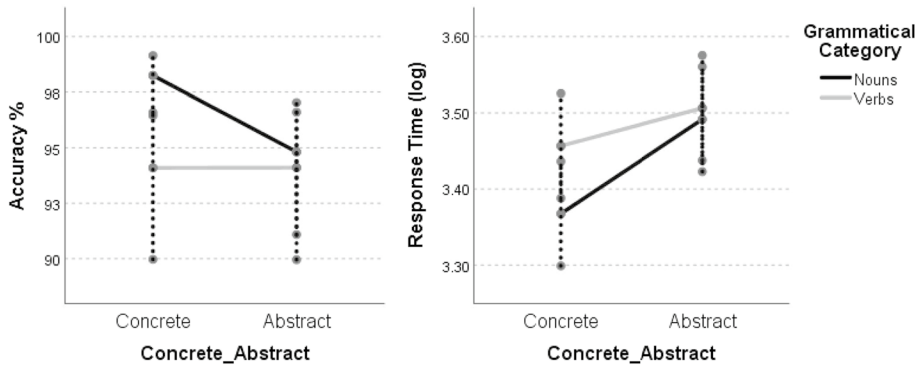


Fig. 1 Mean accuracy and response time across categories

on *highly abstract* noun versus *highly abstract* verb triads for accuracy ( $t(1567) = -.920, p = .856$ )—it is for RT ( $t(1474) = -3.223, p = .014$ ).

Restricting the sample to nouns only, we compared the mean performance of the three concrete noun categories together to the different levels of abstractness in the noun categories. Performance on concrete noun triads was more accurate than that on the *highly abstract* noun triads ( $t(1590) = 2.796, p = .031$ ), but not different from *moderately abstract*



**Fig. 2** Interaction effect between grammatical category and concreteness (error bars represent model-based 95% confidence intervals)

noun triads ( $t(1590)=1.412, p=.498$ ), or *mildly abstract* noun triads ( $t(1590)=.550, p=.704$ ). RT performance was quicker on concrete noun triads than any abstract noun triads ( $p<.001$ ). Restricting the sample to verbs only, we compared the mean performance on the three concrete verb categories together to different levels of abstractness in the verb categories. Concrete verb triads were not judged more accurately than *highly abstract* verb ones ( $t(1568)=2.265, p=.091$ ), *moderately abstract* ( $t(1568)=-1.793, p=.162$ ) or *mildly abstract* ones ( $t(1568)=-1.903, p=.162$ ). Concrete verb triads were judged more quickly than highly abstract verb triads ( $t(1488)=-7.559, p<.001$ ), but not differently from moderately abstract ( $t(1488)=.396, p=.692$ ) or mildly abstract verb triads ( $t(1488)=-1.055, p=.536$ ).

## Discussion

We investigated the effects of grammatical class and levels of concreteness on lexical-semantic processing, finding evidence in favor of a unitary semantic space hypothesis, in which the weight of sensory-perceptual and sensory-motor features influences our conceptual processing abilities. We demonstrated main effects of grammatical category (i.e., nouns versus verbs) and concreteness status (i.e., concrete versus abstract) for both accuracy and RT performance, but notably, detailed analyses into subcategories revealed semantic patterns of lexical organization. An interaction effect showed that the influence of grammatical class was only evident for concrete nouns and verbs, not for abstract ones. Moreover, the concreteness effect was substantially more pronounced for nouns than verbs. Comparisons among specific categories of concrete and abstract nouns and verbs further showed that levels of abstractness have a direct influence on lexical processing performance, and that a word's grammatical class does not affect processing across classifications of mildly, moderately, and highly abstract words. Note that findings across accuracy and RT for grammatical class differences and the concreteness effect were remarkably consistent; higher accuracy corresponded to faster RT, such that there was improvement on both dimensions with increasing concreteness, and for nouns compared to verbs.

Our findings thus favor a unitary semantic space theory, in which concepts reflect—at least in part—a combination of sensory-perceptual and sensory-motor features, which

together form a conceptual representation independent of grammatical class. Additionally, the interaction we found between grammatical class and concreteness does not support a division between nouns and verbs in the organization of our mental lexicon as posed by the grammatical class hypothesis, since abstract nouns and verbs display a different relationship to each other from that between concrete nouns and verbs. Moreover, with our well-controlled design, we demonstrated that the concreteness effect applies not only to nouns but also to verbs, an effect that has scarcely been studied in the literature on lexical-semantic processing in cognitively healthy individuals.

We propose that the previously reported grammatical class effect due to a supposed separate organization of nouns and verbs in the brain (e.g., Hillis and Caramazza 1995) in fact reflects the disguised influence of sensory-perceptual and sensory-motor features on lexical-semantic processing. Concreteness can be viewed as the collective weight of sensory-perceptual and sensory-motor features of a concept (Vinson et al. 2003); indeed, this is what the instructions requested when concreteness ratings were collected by Brysbaert et al. (2014). On average, concrete nouns, often referring to objects that are strongly associated with multiple sensory-perceptual and sensory-motor features, have higher featural weights than concrete verbs (Vinson et al. 2003), which refer to actions that are relatively restricted to sensory-motor features (e.g., Janczyk and Kunde 2012). Consistent with this observation, our results showed that performance was more accurate and quicker on the triads composed of three concrete nouns than that for those composed of three concrete verbs.

Both concrete noun and verb categories were processed more accurately and quickly than abstract noun and abstract verb categories, respectively. Abstract words have fewer sensory-perceptual and sensory-motor features than concrete ones, along a gradient of concreteness. This finding was demonstrated in the main effects, as well as when the data were considered in their narrower categories; within levels of abstractness (i.e., *mildly*, *moderately*, *highly abstract*), results showed a stepwise decrease of performance parallel to the decrease in rated concreteness. Importantly, performance on the triads from the three abstract noun categories was not more accurate than that on the three abstract verb categories, neither in the overall analysis nor when pairs of noun and verb categories with equal levels of abstractness (e.g., *highly abstract* nouns versus *highly abstract* verbs) were compared.

These results do not support the grammatical class hypothesis for lexical organization, as a separation of nouns and verbs in the brain would predict behavior on concrete nouns versus verbs to parallel that on abstract nouns versus verbs. The idea for the grammatical class hypothesis originated from lesion studies in which individuals with post-stroke aphasia performed markedly differently on nouns and verbs (e.g., Shapiro et al. 2000). Subsequent functional neuroimaging studies investigating the predictions of the grammatical class hypothesis have suggested that there are neural distinctions between representations of nouns and verbs (for a review see Crepaldi et al. 2011). However, stimuli used in these studies were primarily concrete words and thus did not test the difference in abstract nouns versus abstract verbs, which only provides one side of the story as the results in the current study indicate.

Instead, our findings are in line with the predictions made by the theory of a unitary semantic space for lexicon (Vigliocco et al. 2004), in which our concepts are organized based on semantically meaningful features, such as sensory, motor, linguistic, and affective information. We propose that the concreteness effect is facilitated by the richer sensory-perceptual representation of concrete words, following Hoffman (2016), and that the featural weight of concrete nouns is greater than that of concrete verbs, following Vinson et al.

(2003). We anticipated that the concreteness effect should be less pronounced in verbs than in nouns, as the relative difference in featural weight is smaller between concrete verbs and abstract verbs than between concrete nouns and abstract nouns. This pattern is indeed evident in our data, as well as in previous reports by Zhang et al. (2006) and Kellenbach et al. (2002), among others. Thus, better performance on concrete nouns compared to concrete verbs and the concreteness effect more generally may well be related to the same underlying mechanism, namely the beneficial influence of the weight of sensory-perceptual and sensory-motor features on semantic processing.

Within the discussion concerning the extent to which grammatical class is an organizational principle in the brain, studies that investigate the interaction between grammatical and semantic factors in neurophysiological and neuroimaging experiments are of particular importance. Barber et al. (2010) carefully manipulated their materials to compare grammatical class versus sensory events in an event-related brain potentials (ERP) experiment and found that grammatical class and semantic effects were virtually identical in latency, duration, and scalp distribution (i.e., topography). Similarly, an ERP study by Pulvermüller et al. (1999) found topographical differences between nouns with strong visual versus action associations (i.e., a semantic contrast), but not between action nouns and action verbs (i.e., a grammatical contrast). In a functional MRI (fMRI) study, Siri et al. (2007) found no verb-specific activation when participants were asked to name the same picture with either an action verb (e.g., to jump) or an action noun (e.g., the jump). Investigating independent contributions of semantic features and grammatical class in a positron emission tomography (PET) study, Vigliocco et al. (2006) found no effects of grammatical class while they were able to distinguish activation differences between sensory and motor features of semantic concepts. Most similar to our behavioral findings, Moseley and Pulvermüller (2014) showed that fMRI activation differed between concrete nouns and verbs, but not between abstract nouns and verbs. Their result emphasizes that the effect between concrete nouns and verbs may be caused by differences in specific sensory-perceptual and sensory-motor features along with overall featural weight, and that this effect disappears in the near absence of such features in abstract words. In sum, these neuroimaging studies strengthen the proposal that behavioral differences between nouns and verbs are driven by semantic factors as opposed to grammatical class.

Our results thus replicated previous reports of the robust concreteness effect that has repeatedly been found in studies of nouns (e.g., James 1975). It is worth noting that many of those reports demonstrate the concreteness effect by contrasting performance on highly concrete words with performance on highly abstract words, using concreteness as a dichotomous variable (e.g., Adorni and Proverbio 2012). The current study went beyond the notion of merely abstract versus concrete and showed a finer-grained, graded effect of concreteness, in which both accuracy and RT performance declined the more abstract the triads became. Remarkably, the concreteness effect was not only observable in RT performance, but also in accuracy performance—typically, accuracy performance of cognitively normal individuals fails to demonstrate the concreteness effect due to a ceiling effect. This finding may be partly explained by our inclusion of individuals aged 50–68 in contrast to the general focus on college students when ‘healthy adults’ are studied. Increasing age, it is known, can result in mild lexical retrieval difficulties (e.g., Barresi et al. 2000). Nonetheless, the fact that the patterns revealed in the current study were consistent across concrete/abstract and noun/verb subcategories (e.g., a graded accuracy decrease as words become more abstract in both nouns and verbs) reinforces our proposed interpretation.

While previous work shows that the sensory-perceptual weight of individual words is strongly related to their concreteness ratings (Buzzeo et al. 2017; Vinson et al. 2003), the

idea that a processing difference between concrete nouns and concrete verbs is caused by a difference in the specific kind and the degree of salience of sensory-perceptual and sensory-motor features should be further explored with future research. A study by Kousta et al. (2011) suggested that while the measures of imageability and concreteness are generally considered to be interchangeable (e.g., Reilly and Kean 2007), imageability may capture additional experiential information, such as emotional valence, better than concreteness does. The authors showed that when holding imageability values constant, concepts with high emotional valence (e.g., peace), which were more abstract than concrete, were processed faster than those with low emotional valence (e.g., menu). Thus, future work should include not only the mean featural weight of semantic categories, but also the proportion of specific kind of features within each concept and, as a result, the weight of individual concepts within a category. The challenge in doing so will be to devise a task assessing single word semantic processing that is sensitive enough to circumvent ceiling effects.

To avoid such ceiling effects, one needs a challenging task that taps into deep semantic processing. Following Sabsevitz et al. (2005), our within-category semantic similarity judgment task was found to satisfy this aim, permitting us to observe subtle patterns in the accuracy and time of lexical-semantic processing. Additionally, the verbal semantic association task also allowed us to test semantic processing of abstract words, which is often hard to induce via, for example, picture naming tasks. Note that a problem with this task is that, because each triad consists of three words, it is not possible to investigate the effect of features of individual words, such as the precise semantic weight of features within a concept as well as psycholinguistic features such as lexical frequency, which have been shown to influence lexical-semantic processing (e.g., Kremin et al. 2001; Vonk 2017; Vonk et al. 2018). Future research should aim to investigate the role of various psycholinguistic features within the framework of concepts' featural weight, to reveal the interaction between semantic and lexical factors in conceptual processing.

We demonstrated that levels of abstractness—as opposed to a dichotomous concrete-abstract classification—matter for lexical processing, since worse accuracy was seen for concrete items with less sensory-perceptual and sensory-motor information than for concrete items with more sensory-perceptual and sensory-motor information. With the idea that the concreteness effect and grammatical class may reflect the same underlying mechanism, namely the weight of sensory-perceptual features, this study provides ground for re-evaluation of previously reported category-specific and grammatical-class effects in neurologically impaired populations, as these phenomena once prompted the idea that nouns and verbs may be separately organized in our mind and brain.

## Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Adorni, R., & Proverbio, A. M. (2012). The neural manifestation of the word concreteness effect: An electrical neuroimaging study. *Neuropsychologia*, *50*(5), 880–891.
- Bak, T. H., & Hodges, J. R. (2003). Kissing and dancing—A test to distinguish the lexical and conceptual contributions to noun/verb and action/object dissociation. Preliminary results in patients with fronto-temporal dementia. *Journal of Neurolinguistics*, *16*(2), 169–181.

- Barber, H. A., Kousta, S., Otten, L. J., & Vigliocco, G. (2010). Event-related potentials to event-related words: Grammatical class and semantic attributes in the representation of knowledge. *Brain Research*, *1332*, 65–74.
- Barresi, B. A., Nicholas, M., Tabor Connor, L., Obler, L. K., & Albert, M. L. (2000). Semantic degradation and lexical access in age-related naming failures. *Aging, Neuropsychology, and Cognition*, *7*(3), 169–178.
- Berndt, R. S., Haendiges, A., Burton, M., & Mitchum, C. (2002). Grammatical class and imageability in aphasic word production: Their effects are independent. *Journal of Neurolinguistics*, *15*(3–5), 353–371.
- Bredin, S. D., Saffran, E. M., & Schwartz, M. F. (1998). Semantic factors in verb retrieval: An effect of complexity. *Brain and Language*, *63*(1), 1–31.
- Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, *150*, 80–84.
- Buzzeo, N. M., Connelly, E., Obler, L. K., Saloner, R., & Vonk, J. M. J. (2017). *The weight of sensory-perceptual features in lexical processing and conceptual imageability*. Poster presented at the Annual Graduate Student Research Poster Day of the Hunter College Communication Sciences and Disorders Program, New York, NY.
- Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, *349*(6312), 788–790. <https://doi.org/10.1038/349788a0>.
- Corp, I. B. M. (2015). *IBM SPSS statistics for windows, version 23*. Armonk, NY: IBM Corp.
- Crepaldi, D., Berlinger, M., Paulesu, E., & Luzzatti, C. (2011). A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effects. *Brain and Language*, *116*(1), 33–49.
- Croft, W. (2000). Parts of speech as language universals and as language-particular categories. In B. C. P. M. Vogel (Ed.), *Approaches to the typology of word classes* (pp. 65–102). New York: Mouton de Gruyter.
- Crutch, S. J., & Warrington, E. K. (2004). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615–627.
- Daniele, A., Giustolisi, L., Silveri, M. C., Colosimo, C., & Gainotti, G. (1994). Evidence for a possible neuroanatomical basis for lexical processing of nouns and verbs. *Neuropsychologia*, *32*(11), 1325–1341.
- Hillis, A. E., & Caramazza, A. (1995). Representation of grammatical categories of words in the brain. *Journal of Cognitive Neuroscience*, *7*(3), 396–407.
- Hoffman, P. (2016). The meaning of ‘life’ and other abstract words: Insights from neuropsychology. *Journal of Neuropsychology*, *10*(2), 317–343.
- Howard, D., & Patterson, K. (1992). *Pyramids and palm trees: A test of semantic access from pictures and words*. Bury St Edmonds: Thames Valley Test Company.
- James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(4), 130–136.
- Janczyk, M., & Kunde, W. (2012). Visual processing for action resists similarity of relevant and irrelevant object features. *Psychonomic Bulletin & Review*, *19*(3), 412–417.
- Jonkers, R. (1998). *Comprehension and production of verbs in aphasic speakers (Dissertation)*. Groningen: Grodil.
- Kellenbach, M. L., Wijers, A. A., Hovius, M., Mulder, J., & Mulder, G. (2002). Neural differentiation of lexico-syntactic categories or semantic features? Event-related potential evidence for both. *Journal of Cognitive Neuroscience*, *14*(4), 561–577.
- Kemmerer, D. (2014). Word classes in the brain: Implications of linguistic typology for cognitive neuroscience. *Cortex*, *58*, 27–51.
- Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., & Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: Evidence from fMRI. *Brain and Language*, *107*(1), 16–43.
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650.
- Kousta, S., 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*(1), 14–34.
- Kremin, H., Perrier, D., De Wilde, M., Dordain, M., Le Bayon, A., Gatignol, P., et al. (2001). Factors predicting success in picture naming in Alzheimer’s disease and primary progressive aphasia. *Brain and Cognition*, *46*(1), 180–183.
- Lee, C., & Federmeier, K. D. (2008). To watch, to see, and to differ: An event-related potential study of concreteness effects as a function of word class and lexical ambiguity. *Brain and Language*, *104*(2), 145–158.

- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.
- Max Planck Institute for Psycholinguistics. (2001). *WebCelex*. Retrieved from <http://celex.mpi.nl/>.
- Moseley, R. L., & Pulvermüller, F. (2014). Nouns, verbs, objects, actions, and abstractions: Local fMRI activity indexes semantics, not lexical categories. *Brain and Language*, 132, 28–42.
- Pulvermüller, F., Mohr, B., & Schleicher, H. (1999). Semantic or lexico-syntactic factors: What determines word-class specific activity in the human brain? *Neuroscience Letters*, 275(2), 81–84.
- Reilly, J., & Kean, J. (2007). Formal distinctiveness of high-and low-imageability nouns: Analyses and theoretical implications. *Cognitive Science*, 31(1), 157–168.
- Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *NeuroImage*, 27(1), 188–200.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.
- Shapiro, K., Shelton, J., & Caramazza, A. (2000). Grammatical class in lexical production and morphological processing: Evidence from a case of fluent aphasia. *Cognitive Neuropsychology*, 17(8), 665–682.
- Silveri, M. C., & Di Betta, A. M. (1997). Noun-verb dissociations in brain-damaged patients: Further evidence. *Neurocase*, 3(6), 477–488.
- Siri, S., Tettamanti, M., Cappa, S. F., Rosa, P. D., Saccuman, C., Scifo, P., et al. (2007). The neural substrate of naming events: Effects of processing demands but not of grammatical class. *Cerebral Cortex*, 18(1), 171–177.
- Tranel, D., Martin, C., Damasio, H., Grabowski, T. J., & Hichwa, R. (2005). Effects of noun-verb homonymy on the neural correlates of naming concrete entities and actions. *Brain and Language*, 92(3), 288–299.
- Van Loon-Vervoorn, W. A. (1985). *Voorstelbaarheidswaarden van Nederlandse woorden: 4600 substantieven, 1000 verba en 500 adjectieven*. Lisse, NL: Swets & Zeitlinger.
- Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience and Biobehavioral Reviews*, 35(3), 407–426.
- 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(4), 422–488.
- Vigliocco, G., Warren, J., Siri, S., Arciuli, J., Scott, S., & Wise, R. (2006). The role of semantics and grammatical class in the neural representation of words. *Cerebral Cortex*, 16(12), 1790–1796.
- Vinson, D. P., & Vigliocco, G. (2002). A semantic analysis of grammatical class impairments: Semantic representations of object nouns, action nouns and action verbs. *Journal of Neurolinguistics*, 15(3), 317–351.
- 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(3), 347–365.
- Vonk, J. M. J. (2015). The contribution of semantic features to the white matter pathways of tool processing. *The Journal of Neuroscience*, 35(37), 12609–12611. <https://doi.org/10.1523/jneurosci.2453-15.2015>.
- Vonk, J. M. J. (2017). *Cognitive and neurobiological degeneration of the mental lexicon in primary progressive aphasia*. (Unpublished doctoral dissertation). The Graduate Center of the City University of New York, New York, NY.
- Vonk, J. M. J., Flores, R. J., Rosado, D., Qian, C., Cabo, R., Habegger, J., et al. (2018). Semantic network function captured by word frequency in nondemented APOE ε4 carriers. *Neuropsychology*. <https://doi.org/10.1037/neu0000508>.
- Warburton, E., Wise, R. J., Price, C. J., Weiller, C., Hadar, U., Ramsay, S., et al. (1996). Noun and verb retrieval by normal subjects: Studies with PET. *Brain*, 119(1), 159–179.
- Zhang, Q., Guo, C., Ding, J., & Wang, Z. (2006). Concreteness effects in the processing of Chinese words. *Brain and Language*, 96(1), 59–68.