Pupillary Responses to Words That Convey a Sense of Brightness or Darkness

Sebastiaan Mathôt, Jonathan Grainger, Kristof Strijkers

Pupillary Responses to Words That Convey a Sense of Brightness or Darkness

Sebastiaan Mathôt[1,2], Jonathan Grainger[2], and Kristof Strijkers[3]
[1]Department of Experimental Psychology, University of Groningen; [2]Laboratoire de Psychologie Cognitive, Unité Mixte de Recherche (UMR) 7290, Centre National de la Recherche Scientifique (CNRS), Aix-Marseille University; and [3]Laboratoire Parole et Langage, UMR 7309, CNRS, Aix-Marseille University

Abstract

Theories about embodiment of language hold that when you process a word's meaning, you automatically simulate associated sensory input (e.g., perception of brightness when you process *lamp*) and prepare associated actions (e.g., finger movements when you process *typing*). To test this latter prediction, we measured pupillary responses to single words that conveyed a sense of brightness (e.g., *day*) or darkness (e.g., *night*) or were neutral (e.g., *house*). We found that pupils were largest for words conveying darkness, of intermediate size for neutral words, and smallest for words conveying brightness. This pattern was found for both visually presented and spoken words, which suggests that it was due to the words' meanings, rather than to visual or auditory properties of the stimuli. Our findings suggest that word meaning is sufficient to trigger a pupillary response, even when this response is not imposed by the experimental task, and even when this response is beyond voluntary control.

Theories about embodiment of language hold that when people process a word's meaning—at least for words that refer to concrete actions or objects—they mentally simulate what they can do with the word's referent and what this referent looks, smells, and feels like. For example, according to such theories, when people read the word *keyboard*, they mentally simulate a typing action, and when they read the word *sun*, they simulate the perception of a bright ball of fire in the sky. Theories positing strong embodiment of language hold that such simulations are necessary for comprehension; that is, to understand what *sun* means, people would need a sensory representation of what it looks like (Glenberg & Gallese, 2012; Pulvermüller, 2013). By contrast, theories suggesting weak embodiment of language are a middle ground between strong embodiment and the traditional view of language as a purely symbolic system that does not involve sensory and motor representations: According to theories suggesting weak embodiment of language, simulations may facilitate language comprehension but are not strictly necessary; that is, mentally picturing the sun may

help someone to read *sun* faster, but they could understand *sun*, even without any sensory representation of it, by relying on a symbolic system (Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012; Zwaan, 2014). As critics of embodied cognition have pointed out, by admitting that sensory simulations may not be necessary for language comprehension, weakly embodied views of language are fundamentally different from strongly embodied views and, in some ways, are similar to traditional, symbolic views of language (Bedny & Caramazza, 2011; Mahon, 2015; Mahon & Caramazza, 2008).

Most support for sensory and motor simulations during language comprehension comes from studies that have taken one of two general approaches: behavioral studies that look at compatibility effects between word

Corresponding Author:
Sebastiaan Mathôt, Department of Experimental Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands
E-mail: s.mathot@cogsci.nl

meaning and action (or perception; Aravena et al., 2012; Glenberg & Kaschak, 2002; Kaschak et al., 2005; Meteyard, Bahrami, & Vigliocco, 2007; Meteyard, Zokaei, Bahrami, & Vigliocco, 2008; Zwaan, Madden, Yaxley, & Aveyard, 2004; Zwaan & Taylor, 2006) and neuroimaging studies that compare brain activity during language comprehension with brain activity during action (or perception; Dravida, Saxe, & Bedny, 2013; Hauk, Johnsrude, & Pulvermuller, 2004; Revill, Aslin, Tanenhaus, & Bavelier, 2008). A compelling example of a behavioral-compatibility effect was reported by Meteyard and her colleagues (2008), who found that upward-downward visual motion affects comprehension speed of words that have a meaning related to upward-downward motion (see also Kaschak et al., 2005; Meteyard et al., 2007); that is, participants decided more quickly that *fall* was a real word (as opposed to a nonword) when they simultaneously saw downward-moving dots. From this, Meteyard et al. (2008) concluded that understanding words that have a meaning related to upward-downward motion involves, at least in part, the same brain areas as perceiving downward motion. This conclusion is supported by neuroimaging studies that show overlap in the brain areas that are active during both (a) reading of words associated with motion and (b) perception of motion (Revill et al., 2008; but for the limits of this overlap, see Dravida et al., 2013).

However, behavioral studies have so far not directly tested one of the central predictions of embodied language: that word meaning by itself can trigger, at least in some cases, associated involuntary actions. For example, consider a landmark study by Glenberg and Kaschak (2002) in which participants judged whether sentences were sensible or not. These sentences conveyed a movement toward or away from the body (e.g., "open the drawer" for movement toward and "close the drawer" for movement away from the body), and participants responded by moving their hands either toward or away from their bodies (e.g., toward themselves for sensible sentences and away from themselves for nonsensible sentences, or the other way around). The crucial finding was that responses were fastest when the direction of the response coincided with the movement direction implied by the sentence; that is, when participants read "open the drawer," they were fastest when they responded with a movement toward the body. This showed that word meaning can modulate actions. However, in this experiment, word meaning did not trigger actions; rather, word meaning sped up (or slowed down) actions that were imposed by the task. To our knowledge, the same is true for all behavioral studies that have investigated the effect of word meaning on action. These studies demonstrate (sometimes very convincingly) that word meaning can modulate actions (e.g., grip force, Aravena et al., 2012) or speed up manual responses (Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006), but not that word meaning can by itself trigger an action.

In the current study, we aimed to show that word meaning by itself can trigger a pupillary light response, an involuntary movement that has traditionally been believed to be a low-level reflex to light. However, recent studies have shown that the light response, although beyond direct voluntary control, is sensitive to high-level cognition (reviewed in Binda & Murray, 2014; Mathôt & Van der Stigchel, 2015). For example, pupils constrict when people covertly attend (without looking at it) to a bright object compared with a dark object (Binda, Pereverzeva, & Murray, 2013; Mathôt, van der Linden, Grainger, & Vitu, 2013; Naber, Alvarez, & Nakayama, 2013). Likewise, pupils constrict when people imagine a bright object (Laeng & Sulutvedt, 2014) or when a bright object reaches awareness in a binocular-rivalry paradigm (Naber, Frassle, & Einhauser, 2011). These phenomena are often explained in terms of top-down modulation of visual brain areas (Mathôt, Dalmaijer, Grainger, & Van der Stigchel, 2014; Wang & Munoz, 2016); that is, pupils constrict when people covertly attend to a bright object, because attention enhances the representation of the bright object throughout visual cortical and subcortical areas.

This reasoning can be naturally extended to embodied language: If word comprehension activates brain areas known to be involved in processing of nonlinguistic visual information (i.e., creates sensory representations), then understanding words that convey a sense of brightness or darkness should trigger pupillary responses—just like attending to (Mathôt et al., 2013) or imagining (Laeng & Sulutvedt, 2014) bright or dark objects. Phrased differently, if words that convey brightness trigger a pupillary constriction relative to words that convey darkness, this would support the view that word comprehension affects sensory brain areas and can even trigger involuntary (pupillary) responses. To our knowledge, this would also be the first direct demonstration that word comprehension by itself can trigger a response and not merely modulate an action that has been imposed by task instructions.

To test this hypothesis, we conducted two main experiments in which participants read (visual experiment) or listened to (auditory experiment) words conveying brightness or darkness. We measured pupil size and predicted that pupils would be smaller when participants read or listened to words conveying brightness than when they read or heard words conveying darkness: a pupillary light response triggered by word meaning.

Method

Stimuli

For the main experiments, in which we varied the *semantic brightness* of words (i.e., whether words conveyed brightness or darkness), we manually selected 121 words from Lexique (New, Pallier, Brysbaert, & Ferrand, 2004), a large database with lexical properties of French words.

There were four word categories: words conveying brightness (e.g., *illuminé* or "illuminated"; *n* = 33), words conveying darkness (e.g., *foncé* or "dark"; *n* = 33), neutral words (e.g., *renforcer* or "to reinforce"; *n* = 35), and animal names (e.g., *lapin* or "rabbit"; *n* = 20). During the visual experiment, words were shown as dark letters (8.5 cd/m^2) against a gray background (17.4 cd/m^2). For the auditory experiment, words were generated in a synthetic voice by using the Mac OS X "say" command to convert text to speech.

Because we wanted to compare pupillary responses to words conveying brightness or darkness, these two categories needed to be matched as accurately as possible. We focused on two main properties: lexical frequency, or how often a word occurs in books (words conveying brightness: *M* = 41 per million, *SD* = 147; words conveying darkness: *M* = 39 per million, *SD* = 119), and, for the visual experiment, visual intensity (words conveying brightness: *M* = 1.58 × 10^6 arbitrary units, *SD* = 4.31 × 10^5; words conveying darkness: *M* = 1.56 × 10^6 arbitrary units, *SD* = 4.26 × 10^5). Visual intensity was matched by selecting words that had approximately the same number of letters, then generating images of these words, and finally iteratively resizing these images until the visual intensity (i.e., summed luminance) of the words was almost identical between the two categories.

In the end, we had a stimulus set in which words conveying brightness or darkness were very closely matched; however, as a result of our stringent criteria, our set contained several variations of the same words, such as *briller* ("to shine") and *brillant* ("shining"). But given pupils' sensitivity to slight differences in task difficulty (i.e., lexical frequency) and visual intensity, we felt that matching was more important than having a highly varied stimulus set.

For the control experiment, in which we varied the valence of words, we selected 60 words that were rated for valence by Bonin et al. (2003), complemented with the 20 animal names selected for the main experiments. Positive words (e.g., *cadeau* or "present"; *n* = 30) had a valence of at least 3.5 on a scale from 1 (*negative*) to 5 (*positive*), and negative words (e.g., *cicatrice* or "scar"; *n* = 30) had a valence of 2.5 or less. The positive and negative words were matched on lexical frequency (positive: *M* = 3.21, *SD* = 0.76; negative: *M* = 3.26, *SD* = 0.53) and visual intensity (positive: *M* = 1.15 × 10^6, *SD* = 3.47 × 10^5; negative: *M* = 1.15 × 10^6, *SD* = 3.48 × 10^5), and none of the words had any obvious association with brightness or darkness.

For all experiments, stimuli were manually selected on the basis of strict criteria. Our sample size of around 30 words per condition was therefore a compromise between having well-matched stimuli and having a reasonable number of observations per participant and condition.

Pupillometry experiments

Thirty naive observers (age range: 18–54 years; 21 women, 9 men) participated in the visual experiment. Thirty other naive observers participated in the auditory experiment (age range: 18–31 years; 19 women, 11 men). Finally, 30 naive observers participated in the control experiment, four of whom had also participated in the auditory experiment (age range: 18–31 years; 19 women, 11 men). We used two to three times as many participants per experiment as in most previous studies on the pupillary light response (e.g., *n* = 5–15 participants in Binda et al., 2013; Mathôt et al., 2013, but 52 participants in Experiment 5 of Laeng & Sulutvedt, 2014), because we expected the effect of embodied language on pupil size to be relatively small. Participants reported normal or corrected vision, provided written informed consent before the experiment, and received €6 for their participation. The experiment was conducted with approval of the Comité d'éthique de l'Université d'Aix-Marseille (Ref. 2014–12–03–09).

Pupil size was recorded monocularly with an EyeLink 1000 (SR Research, Mississauga, ON, Canada), a video-based eye tracker sampling at 1000 Hz. Stimuli were presented on a 21-in. CRT monitor (screen resolution: 1,024 × 768 pixels; refresh rate: 150 Hz; model p227f, ViewSonic, Walnut, CA). Testing took place in a dimly lit room. The experiment was implemented with OpenSesame (Mathôt, Schreij, & Theeuwes, 2012) using the Expyriment (Krause & Lindemann, 2014) back end.

At the beginning of each session, a nine-point eye-tracker calibration was performed. Before each trial, a single-point recalibration (drift correction) was performed. Each trial started with a dark central fixation dot on a gray background for 3 s. Next, a word was presented. In the visual experiment and the control experiment, the word was presented in the center of the screen for 3 s or until the participant pressed the space bar; in the auditory experiment, the word was played back through a set of desktop speakers, and the experiment paused for 3 s or until the participants pressed the space bar. The participants' task was to press the space bar whenever they saw or heard an animal name and to withhold response otherwise. Participants saw or heard each word once, with the exception of *pénombre* in the visual experiment.[1] Word order was fully randomized.

Normative ratings

For all words conveying brightness or darkness, we collected normative ratings from 30 naive observers (age range: 18–29 years; 17 women, 13 men), most of whom had not participated in the pupillometry experiments. Participants received €2 for their participation. Words

were presented one at a time and using the same images used for the visual pupillometry experiment, together with a five-point rating scale. On this scale, participants indicated how strongly the word conveyed a sense of brightness (from *very bright* to *very dark*) or the word's valence (from *very negative* to *very positive*). Brightness and valence were rated in separate blocks, the order of which was counterbalanced across participants. On the basis of valence ratings, we calculated the emotional intensity of the words as the deviation from neutral valence (intensity = $|3 - \text{valence}|$).

Results

Trials in which participants' responses were incorrect (i.e., false alarms and misses) were excluded (visual experiment: 0.36% of trials; auditory experiment: 0.44% of trials; control experiment: 1.12% of trials), as were trials in which the pupil-size recordings contained artifacts (i.e., trials in which pupil size was less than 0.5 times or more than 2.5 times baseline size, because this generally corresponded to unreconstructed blinks or signal loss; visual experiment: 10.1% of trials; auditory experiment: 4.5% of trials; control experiment: 4.2% of trials). The average response time to animal words was 790 ms (*SE* = 8.6) for the visual experiment, 1067 ms (*SE* = 12.5) for the auditory experiment, and 790 ms (*SE* = 10.0) for the control experiment. No participants were excluded from the analysis.

Pupil-size results

The main results (for the visual and auditory experiments) are shown in Figure 1, in which pupil size is plotted over time from word onset, separately for words conveying brightness, words conveying darkness, and neutral words. (Animal names are not shown, because the pupillary response was distorted by participants' key-press responses.) As predicted, pupils were smaller when participants read or heard words conveying brightness compared with words conveying darkness. This effect was present for both visually presented words (Fig. 1a) and spoken words (Fig. 1b). The effect arose gradually and slowly, and it peaked between 1 and 2 s after word onset. For neutral words, which did not convey a specific sense of brightness, pupil size was intermediate.

In addition to the effect of semantic brightness, in the visual experiment, there was a pronounced pupillary dilation that peaked around 0.6 s after word onset (Fig. 1a). This is an alerting effect, or orienting response (Mathôt et al., 2014; Wang & Munoz, 2015). This early pupillary response was not clearly modulated by the semantic brightness of the words and was followed by the pronounced constriction that always follows visual changes (e.g., Mathôt, Melmi, & Castet, 2015). In the auditory experiment, there was no visual stimulation to trigger pupillary constriction, and pupils therefore dilated throughout the trial, with no clear distinction between the early orienting response and later dilation due to task-related effort.

Pupil size is reported as a proportion of pupil size at word onset and was smoothed with a 51-ms Hanning window. Blinks were reconstructed with cubic-spline interpolation (Mathôt, 2013). Only words conveying brightness or darkness were carefully matched (see the Method section), and therefore only these two categories were included in statistical tests. (However, including all
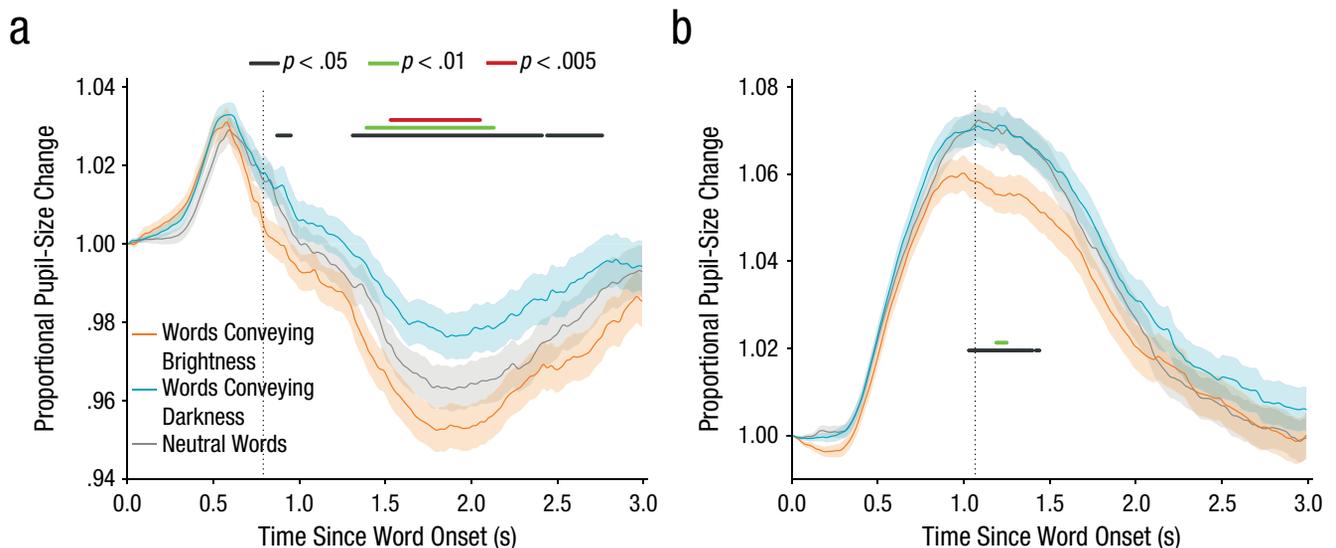

Fig. 1. Proportional change in pupil size as a function of time, separately for words conveying brightness, words conveying darkness, and neutral words. Results are shown for the (a) visual experiment and (b) auditory experiment. The shaded areas represent ±1 *SE*. Horizontal lines indicate periods during which pupil size differed significantly between trials with words conveying brightness and trials with words conveying darkness, using three different α thresholds. The vertical dotted lines indicate the mean response time to animal words.
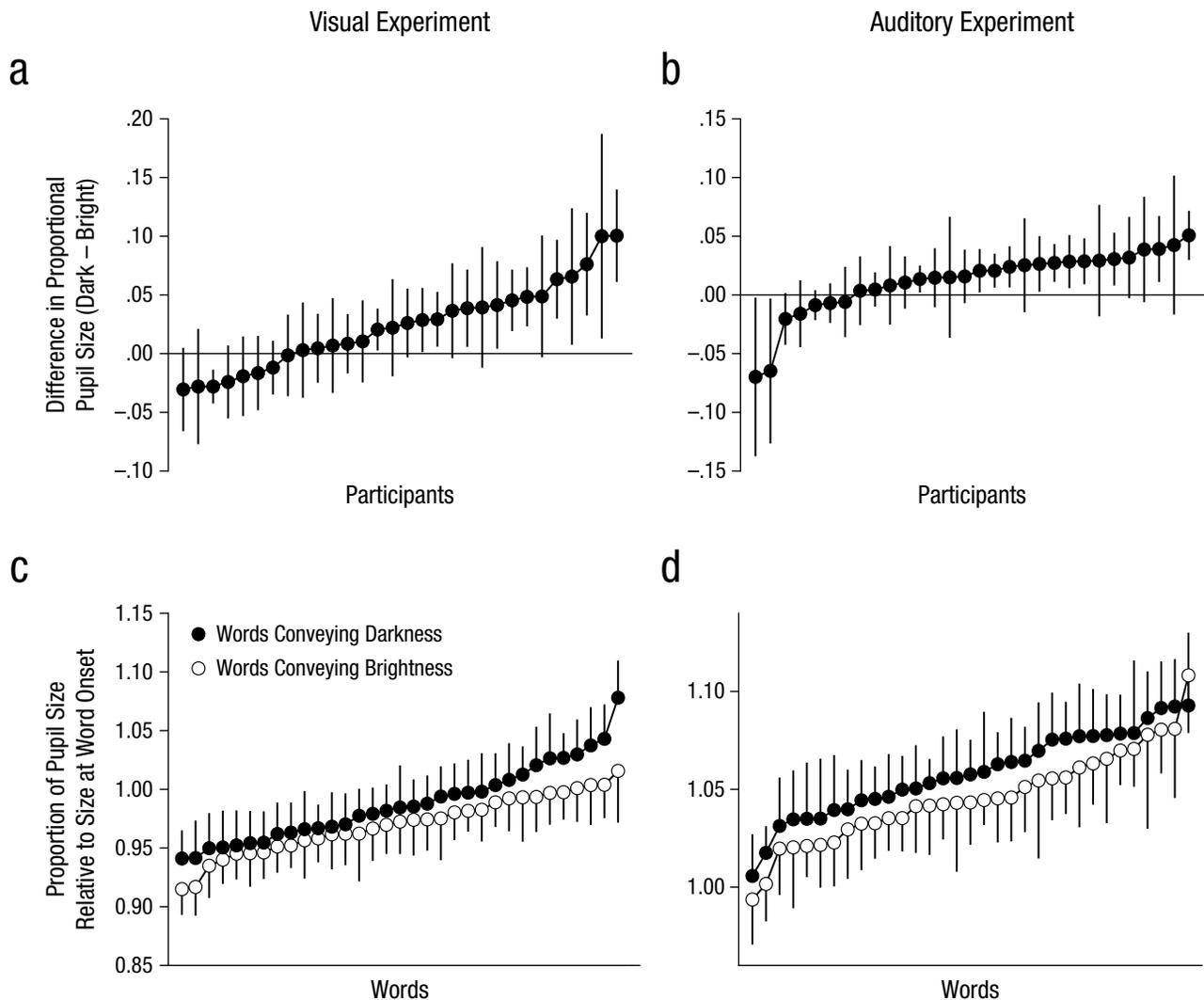
Fig. 2. Results for individual participants (top row) and individual words (bottom row). Mean proportional pupil-size change in the (a) visual experiment and (b) auditory experiment is presented separately for each participant. The data points are ordered by the magnitude of the difference in change in pupil size, which was calculated as pupil size for words conveying darkness minus pupil size for words conveying brightness. Mean proportion of pupil size relative to size at word onset in the (c) visual experiment and (d) auditory experiment is presented separately for each of the words conveying darkness and words conveying brightness. The data points are ordered by the magnitude of the change in pupil size. Error bars indicate ±1 *SE*. Pupil size was measured during the 1- to 2-s window after stimulus onset.

words yielded similar results; see the Supplemental Material available online.) For each 10-ms window, we conducted a linear-mixed effects model using the lme4 package (Version 1.1-12; Bates, Maechler, Bolker, & Walker, 2016) for the R software environment (Version 3.3.2; R Development Core Team, 2016). In this model, pupil size was the dependent variable, and semantic brightness (bright or dark) was the fixed effect; we used random by-participant intercepts and slopes. The R formula for our model was as follows—model: pupil_size ~ brightness + (1 + brightness | participant). We commonly use a significance threshold of at least 200 contiguous milliseconds in which p is less than .05 (Mathôt et al., 2013). To estimate p values, we interpreted t values as though they were z values (i.e., the computationally fast

t-as-z approach), such that a t of 1.96 corresponded to a p of .05. This approach was anticonservative, but only slightly so given our sample size (Luke, 2016). With this criterion, the semantic-brightness effect was reliable from 1,310 to 2,410 ms and from 2,440 to 2,760 ms in the visual experiment and from 1,030 to 1,360 ms in the auditory experiment. However, Figure 1 shows significant differences in pupil size between words conveying brightness and words conveying darkness as determined using three alpha thresholds and no minimum number of contiguous samples.

To test how general the effect was, we also looked at mean pupil size during the 1- to 2-s window for individual participants and words. As shown in Figures 2a and 2b, the majority of the participants showed an effect in

the predicted direction, and this effect was supported by a default Bayesian one-sided, one-sample *t* test, conducted with JASP (Version 0.7; Love et al., 2015). For the visual experiment, the Bayes factor (BF) was 51.3, and for the auditory experiment, the BF was 4.1; the combined BF was 211.1, which represents decisive evidence in support of a positive effect. As shown in Figures 2c and 2d, the effect of semantic brightness was small relative to the variability between words; however, there was a clear shift in the distribution of pupil sizes, so that pupil size was slightly larger for words conveying darkness than for words conveying brightness, which was again supported by a default Bayesian one-sided, independent-samples *t* test. For the visual experiment, the BF was 6.7, and for the auditory experiment, the BF was 3.8; the combined BF was 25.4, which represents strong evidence in support of the hypothesis that pupils are larger in response to words conveying darkness than in response to words conveying brightness. The results were similar when we analyzed the full 3 s of the trials rather than only the 1- to 2-s window (see the Supplemental Material).

The effects of valence and emotional intensity

To test whether the effect of semantic brightness could be due to differences in valence or emotional intensity, we analyzed the subjective ratings for semantic brightness, valence (positive or negative), and emotional intensity of all words.

First, participants rated words on a scale from 1 (*bright*) to 5 (*dark*). They agreed with our own division of the words into a set of words conveying brightness (rated darkness: $M = 0.84$, $SD = 0.2$) and a set of words conveying darkness (rated darkness: $M = 3.00$, $SD = 0.24$), $t(64) = 36.43$, $p < .001$ (independent-samples *t* test). Second, we found a strong and reliable correlation between brightness and valence ($r = .89$, $p < .001$), such that words conveying brightness were rated as being more positive than words conveying darkness. Valence probably has no effect on pupil size beyond that of emotional intensity; that is, pupils dilate similarly in response to negative and positive stimuli of equal emotional intensity (Collins, Ellsworth, & Helmreich, 1967). The correlation between brightness and valence was so strong that we could not control for it statistically. Therefore, we conducted a control experiment in which we looked at the pupillary response to positive and negative words that had no association with brightness. This experiment confirmed that valence has no notable effect on pupil size (see the Supplemental Material): A default Bayesian two-sided one-sample *t* test conducted on a per-participant basis provided substantial evidence for the null hypothesis (BF = 4.9 in support of no effect); a default Bayesian

two-sided independent-samples *t* test conducted on a per-word basis also provided substantial evidence for null hypothesis (BF = 3.0 in support of no effect). (If valence had been a confound in our main experiments, pupils would have been considerably larger in response to negative words than to positive words.) This control experiment also showed that emotional intensity (i.e., deviation from neutral valence) strongly affected pupil size, such that participants' pupils were largest in response to emotionally intense words (from 810 ms until trial end)—regardless of whether words were intensely positive or intensely negative.

Third, we found a weak but reliable correlation between brightness and emotional intensity ($r = .22$, $p = .027$), such that bright words were rated as being more emotionally intense than dark words; given the results from our control experiment (see also Goldwater, 1972), this finding would drive an effect in the opposite direction from the effect that we observed, because emotionally intense stimuli (bright words in our case) trigger a strong pupillary dilation. Given that the correlation between brightness and emotional intensity was only weak, we could take the effect of emotional intensity into account statistically by conducting a control analysis that was identical to the regression analysis described earlier but included emotional intensity as control predictor. Conducted using the same criteria as described earlier, this analysis again revealed in both experiments an effect of brightness that was reliable and in the expected direction. During the visual experiment, the effect was observed between 1,260 and 2,590 ms after word onset; during the auditory experiment, the effect was observed between 830 and 1,460 ms. In the auditory experiment, we also observed an effect between 90 and 330 ms; given that this effect was small and extremely early, it was probably spurious. In summary, it is theoretically and statistically unlikely that the effect of semantic brightness on pupil size was driven by differences in the valence and emotional intensity of our stimuli.

Discussion

We have demonstrated that the eyes' pupils are smaller after people read or listen to words conveying brightness (e.g., *sun*) than when people read or listen to words conveying darkness (e.g., *night*). This effect arises slowly and gradually and, in our experiments, peaked between 1 and 2 s after word onset.

Our findings have important implications for theories about how motor and sensory simulations arise during language comprehension. Our starting premise was that there is a cognitive pupillary response to light (i.e., one without direct visual stimulation) that reflects sensory representations that are similar to those that arise during perception

(Laeng & Sulutvedt, 2014) and involve brain areas engaged in processing of nonlinguistic visual information (Mathôt et al., 2014). Our findings therefore suggest that word comprehension can induce activity in nonlinguistic visual brain areas, and can even trigger involuntary responses (e.g., a pupillary response) that correspond to the word's meaning. This finding is consistent with results of previous behavioral studies showing that word meaning can modulate actions (e.g., Aravena et al., 2012; Glenberg & Kaschak, 2002). However, in previous studies, actions were not triggered by word meaning per se but were imposed by the task and voluntarily performed by the participants (e.g., Aravena et al., 2012; Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006). Our findings extend these studies by showing that word meaning alone is sufficient to trigger a response that is not imposed by the task (i.e., the instructions did not mention pupils) and that is largely beyond voluntary control (see also Spivey & Geng, 2001, who showed that participants had spontaneous eye movements during mental imagery).

Is pupillary constriction merely a nonfunctional by-product of reading words conveying brightness, or does it serve some function? If the latter, what might that function be? Given that the semantic pupillary effect occurs late—in the visual experiment, it did not reach its peak until about a second after participants had processed the word's meaning and responded to it—it is unlikely that pupillary responses themselves played a role in word comprehension. Therefore, our results do not reveal whether word representations are strongly embodied (consist, in part, of sensory and motor simulations) or are not (instead are accompanied by sensory and motor simulations; Mahon, 2015). However, our results do show that word meaning triggers mental simulations and can even trigger physiological (i.e., pupillary) responses— even when these responses are irrelevant to the task, and even when they are largely beyond voluntary control. This shows a profound interaction between language and the sensory and motor systems.

One possible function of sensory and motor simulations during language comprehension—one that is not often considered in discussions of embodied language— is preparation for the immediate future. For example, when people hear the name of an object that is in their field of view, they are likely to look at it (Cooper, 1974). Therefore, the pupillary constriction that is triggered by reading the word *lamp* may reflect preparation for looking at a lamp (Mathôt, van der Linden, Grainger, & Vitu, 2015). In this view, sensory and motor simulations that arise during word comprehension are not (or not only) part of the comprehension process but reflect preparation for actions and perceptions that are likely to occur in the immediate future. There are many ways, direct and

indirect, automatic and voluntary, in which language causes preparation. For example, if someone reads a scary passage from a horror novel about a rabid dog, his or her body might show arousal responses (goose bumps, etc.) that prepare the body for action. If this happens, the link between language and action preparation is automatic, in the sense that goose bumps, like pupillary responses, are involuntary. However, the link is also indirect, because it depends not only on language comprehension but also on fear: People who are afraid of dogs may get goose bumps when they read the novel about the rabid dog, but people who own a big dog may not.

In contrast, in the case of the semantic pupillary response, the link between language and action is direct, dependent only on language comprehension: Visual brain areas are automatically activated when the brain processes visual words (e.g., Revill et al., 2008), and whenever a neural representation of brightness is activated, either through microstimulation of visual brain areas (Wang & Munoz, 2016) or through cognitive processes such as attention (Mathôt et al., 2013) and mental imagery (Laeng & Sulutvedt, 2014), pupils constrict. Demonstrations of direct, automatic sensory and motor activation during language comprehension (as in the present study) could also be interpreted as reflecting a very direct link between language and action preparation. It is easy to see how such direct links could be beneficial (e.g., by adjusting pupil size ahead of time for optimal perception of an object to be looked at), regardless of whether the stuff of thought is fundamentally embodied or not—which is still an open question.

Finally, studies on embodied language have not always proven replicable. A study by Papesh (2015) is particularly relevant: She failed to replicate the findings of Glenberg and Kaschak (2002)—a key study that we have highlighted here as well. To date, there have been no large-scale replications of important embodied-cognition effects and no systematic meta-analyses to estimate how strongly the field is distorted by selective reporting and publication bias. But studies such as that by Papesh suggest caution—although not, in our view, extreme skepticism. Therefore, in the present study, we have provided a complete description of all data, a conceptual replication (i.e., the auditory experiment), and public access (for the URL, see the Open Practices statement at the end of the article) to all relevant materials to allow for independent, direct replication.

In summary, we have shown that pupils constrict when people read or listen to words conveying brightness compared with words conveying darkness. This shows that word comprehension alone is sufficient to activate sensory and motor representations and can even trigger involuntary (pupillary) responses.

Action Editor

Author Contributions

Declaration of Conflicting Interests

Funding

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617702699

Open Practices

All data and materials have been made publicly available via Figshare and can be accessed at https://doi.org/10.6084/m9.figshare.4742935.v2. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617702699. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at https://www.psychologicalscience.org/publications/badges.

Note

1. Because of a bug in the experiment, the word *pénombre* was shown twice; this is why there were slightly more words conveying darkness than words conveying brightness.

References

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., & Nazir, T. (2012). Grip force reveals the context sensitivity of language-induced motor activity during "action words" processing: Evidence from sentential negation. *PLoS ONE*, *7*(12), Article e50287. doi:10.1371/journal.pone.0050287

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2016). lme4: Linear mixed-effects models using 'Eigen' and S4 (Version 1.1-12) [Computer software]. Retrieved from http://cran.r-project.org/package=lme4

Bedny, M., & Caramazza, A. (2011). Perception, action, and word meanings in the human brain: The case from action verbs. *Annals of the New York Academy of Sciences*, *1224*, 81–95. doi:10.1111/j.1749-6632.2011.06013.x

Binda, P., & Murray, S. O. (2014). Keeping a large-pupilled eye on high-level visual processing. *Trends in Cognitive Sciences*, *19*, 1–3. doi:10.1016/j.tics.2014.11.002

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience*, *33*, 2199–2204. doi:10.1523/jneurosci.3440-12.2013

Bonin, P., Méot, A., Aubert, L., Malardier, N., Niedenthal, P., & Capelle-Toczek, M. (2003). Normes de concrétude, de valeur d'imagerie, de fréquence subjective et de valence émotionnelle pour 866 mots [Norms for concreteness, imaging value, subjective frequency and emotional valence for 866 words]. *L'année Psychologique*, *103*, 655–694.

Collins, B. E., Ellsworth, P. C., & Helmreich, R. L. (1967). Correlations between pupil size and the semantic differential: An experimental paradigm and pilot study. *Psychonomic Science*, *9*, 627–628. doi:10.3758/bf03327922

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84–107. doi:10.1016/0010-0285(74)90005-x

Dravida, S., Saxe, R., & Bedny, M. (2013). People can understand descriptions of motion without activating visual motion brain regions. *Frontiers in Psychology*, *4*, Article 537. doi:10.3389/fpsyg.2013.00537

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, *48*, 905–922. doi:10.1016/j.cortex.2011.04.010

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558–565. doi:10.3758/bf03196313

Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin*, *77*, 340–355. doi:10.1037/h0032456

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301–307. doi:10.1016/s0896-6273(03)00838-9

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*, B79–B89. doi:10.1016/j.cognition.2004.06.005

Krause, F., & Lindemann, O. (2014). Expyriment: A python library for cognitive and neuroscientific experiments. *Behavior Research Methods*, *46*, 416–428. doi:10.3758/s13428-013-0390-6

Laeng, B., & Sulutvedt, U. (2014). The eye pupil adjusts to imaginary light. *Psychological Science*, *25*, 188–197. doi:10.1177/0956797613503556

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., . . . Wagenmakers, E. J. (2015). JASP

(Version 0.7) [Computer software]. Retrieved from https://jasp-stats.org/

Luke, S. G. (2016). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*. Advance online publication. doi:10.3758/s13428-016-0809-y

Mahon, B. Z. (2015). What is embodied about cognition? *Language, Cognition and Neuroscience*, 30, 420–429. doi:10.1080/23273798.2014.987791

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, 102, 59–70. doi:10.1016/j.jphysparis.2008.03.004

Mathôt, S. (2013). *A simple way to reconstruct pupil size during eye blinks*. doi:10.6084/m9.figshare.688001

Mathôt, S., Dalmaijer, E., Grainger, J., & Van der Stigchel, S. (2014). The pupillary light response reflects exogenous attention and inhibition of return. *Journal of Vision*, 14(14), Article 7. doi:10.1167/14.14.7

Mathôt, S., Melmi, J. B., & Castet, E. (2015). Intrasaccadic perception triggers pupillary constriction. *PeerJ, 3*, Article e1150. doi:10.7717/peerj.1150

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44, 314–324. doi:10.3758/s13428-011-0168-7

Mathôt, S., van der Linden, L., Grainger, J., & Vitu, F. (2013). The pupillary response to light reflects the focus of covert visual attention. *PLoS ONE*, 8(10), Article e78168. doi:10.1371/journal.pone.0078168

Mathôt, S., van der Linden, L., Grainger, J., & Vitu, F. (2015). The pupillary light response reflects eye-movement preparation. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 28–35. doi:10.1037/a0038653

Mathôt, S., & Van der Stigchel, S. (2015). New light on the mind's eye: The pupillary light response as active vision. *Current Directions in Psychological Science*, 24, 374–378. doi:10.1177/0963721415593725

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science*, 18, 1007–1013. doi:10.1111/j.1467-9280.2007.02016.x

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. doi:10.1016/j.cortex.2010.11.002

Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology*, 18, 732–733. doi:10.1016/j.cub.2008.07.016

Naber, M., Alvarez, G. A., & Nakayama, K. (2013). Tracking the allocation of attention using human pupillary oscillations. *Frontiers in Psychology, 4*, Article 919. doi:10.3389/fpsyg.2013.00919

Naber, M., Frassle, S., & Einhauser, W. (2011). Perceptual rivalry: Reflexes reveal the gradual nature of visual awareness. *PLoS ONE*, 6(6), Article e20910. doi:10.1371/journal.pone.0020910

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36, 516–524. doi:10.3758/bf03195598

Papesh, M. H. (2015). Just out of reach: On the reliability of the action-sentence compatibility effect. *Journal of Experimental Psychology: General*, 144, e116–e141. doi:10.1037/xge0000125

Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, 17, 458–470. doi:10.1016/j.tics.2013.06.004

R Development Core Team. (2016). R: A language and environment for statistical computing (Version 3.3.2) [Computer software]. Retrieved from https://www.r-project.org/index.html

Revill, K. P., Aslin, R. N., Tanenhaus, M. K., & Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proceedings of the National Academy of Sciences, USA*, 105, 13111–13115. doi:10.1073/pnas.0807054105

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, 65, 235–241. doi:10.1016/j.tics.2013.06.004

Wang, C., & Munoz, D. P. (2015). A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology*, 33, 134–140. doi:10.1016/j.conb.2015.03.018

Wang, C., & Munoz, D. P. (2016). Attentional modulation of pupillary light responses by microstimulation of the superior colliculus. *Journal of Vision*, 16, Article 936. doi:10.1167/16.12.936

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences*, 18, 229–234. doi:10.1016/j.tics.2014.02.008

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science*, 28, 611–619. doi:10.1016/j.cogsci.2004.03.004

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1–11. doi:10.1037/0096-3445.135.1.1