AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Review

© 2025 American Psychological Association
ISSN: 0033-295X

https://doi.org/10.1037/rev0000546

Cognitive and Neural Mechanisms of Linguistic Influence on Perception

Ksenija Slivac[1, 2], Peter Hagoort[1, 3], and Monique Flecken[4]
[1] Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
[2] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[3] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen
[4] Department of Linguistics, University of Amsterdam

To date, research has reliably shown that language can engage and modify perceptual processes in a top-down manner. However, our understanding of the cognitive and neural mechanisms underlying such top-down influences is still under debate. In this review, we provide an overview of findings from literature investigating the organization of semantic networks in the brain (spontaneous engagement of the visual system while processing linguistic information), and linguistic cueing studies (looking at the immediate effects of language on the perception of a visual target), in an effort to isolate such mechanisms. Additionally, we connect the findings from linguistic cueing studies to those reported in (nonlinguistic) literature on priors in perception, in order to find commonalities in neural processes allowing for top-down influences on perception. In doing so, we discuss the effects of language on perception in the context of broader, general cognitive and neural principles. Finally, we propose a way forward in the study of linguistic influences on perception.

Keywords: language–perception interaction, semantics, visual perception, neural mechanisms, top-down influences

The language system in the brain is extensively intertwined with other functional systems, such as perception, attention, emotions, and the sensorimotor system (Binder et al., 2009; Huth et al., 2016). In the last few decades, the extent to which language can and does regularly affect these systems, especially higher and lower level perception, and modify neural activation therein, has become of particular interest to cognitive (neuro)scientists (Barsalou, 2008; Borghi & Cimatti, 2009; Dove, 2020; Louwerse, 2011; Tillas, 2015).

The line of research concerned with the ability of language to engage perceptual processes is quite multifaceted. On the one hand, neuroscientists interested in neural correlates of linguistic processing have studied the scope of semantic representations in the brain and the ability of language to engage the perceptual systems in the absence of visual stimuli. While we focus on visual perception in this review, it should be noted that a substantial amount of work has been done on the extent of semantic encoding with respect to sensorimotor cortices (Kiefer & Pulvermüller, 2012; Pulvermüller, 2013). This line of research views language as interacting with our perceptual systems and being capable of stimulating those systems in order to reenact or reexperience what has been expressed through language (Barsalou, 2008). This view acknowledges that language allows us to systematically document, think, and communicate about the world around us, facilitating both retention and transmission of our thoughts and experiences. By forming associations between linguistic and nonlinguistic aspects of our cognition and sensation, we end up using language as a shortcut or tool for conceptualization, removing the necessity to fully experience firsthand everything that has been verbally expressed in order to understand it. In other words, with language, we can simulate our experiential states in a retroactive manner, both behaviorally—by reinforcing or augmenting our understanding of what has been expressed through language, and neurally—by coactivating neural systems that encode such states (Barsalou, 2008; Borghi & Cimatti, 2009; Tillas, 2015).

On the other hand, neuroscientists interested in the linguistic effect on our perceptual experience have investigated the extent to which the perceptual systems are susceptible to top-down influences such as attention, expectation, prior beliefs, and also language (C. D. Gilbert & Li, 2013; C. D. Gilbert & Sigman, 2007). Top-down influences on perception have been studied mainly by employing a cueing paradigm, with a cue, such as a symbol, word, picture, or sound, presented shortly before a visual stimulus (target) upon which participants are asked to perform a task. This approach allows for a highly controlled environment in which researchers can outline

---

Elena L. Grigorenko served as action editor.

Ksenija Slivac https://orcid.org/0000-0003-0806-1898

Figures 1 and 2 have been created using images from https://Freepik.com.

This research was funded in whole, or in part, by the Netherlands Organisation for Scientific Research (Grant 024.001.006, awarded to Peter Hagoort). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Ksenija Slivac played a lead role in conceptualization, visualization, writing–original draft, and writing–review and editing. Peter Hagoort played a supporting role in conceptualization, supervision, writing–original draft, and writing–review and editing. Monique Flecken played a supporting role in conceptualization, supervision, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Ksenija Slivac, Max Planck Institute for Human Cognitive and Brain Sciences, Postfach 500355, D-04303 Leipzig, Germany. Email: ksenija.slivac@mpi.nl or slivac@cbs.mpg.de

how the presence of a preceding stimulus can affect the perception of the visual target and/or the neural pattern of activation in response to that target.

This paradigm has also been adopted by researchers interested in linguistic top-down effects on perception. A series of behavioral and neuroimaging studies have revealed that language seems to be particularly efficient in modulating our perceptual experiences. In particular, several cueing studies have directly compared the effect of linguistic labels (e.g., word "dog") to that of visual (e.g., picture of a dog) or auditory (e.g., barking sound) cues on visual perception, showing that people respond to a visual task more quickly and accurately when cued by linguistic rather than visual or auditory cues (Boutonnet & Lupyan, 2015; Lupyan & Spivey, 2010a). These findings have put forward the notion that linguistic labels exert a stronger influence on perceptual processes than seemingly equivalent environmental sounds or pictures (Edmiston & Lupyan, 2015).

Theoretical Account of Language–Perception Interaction

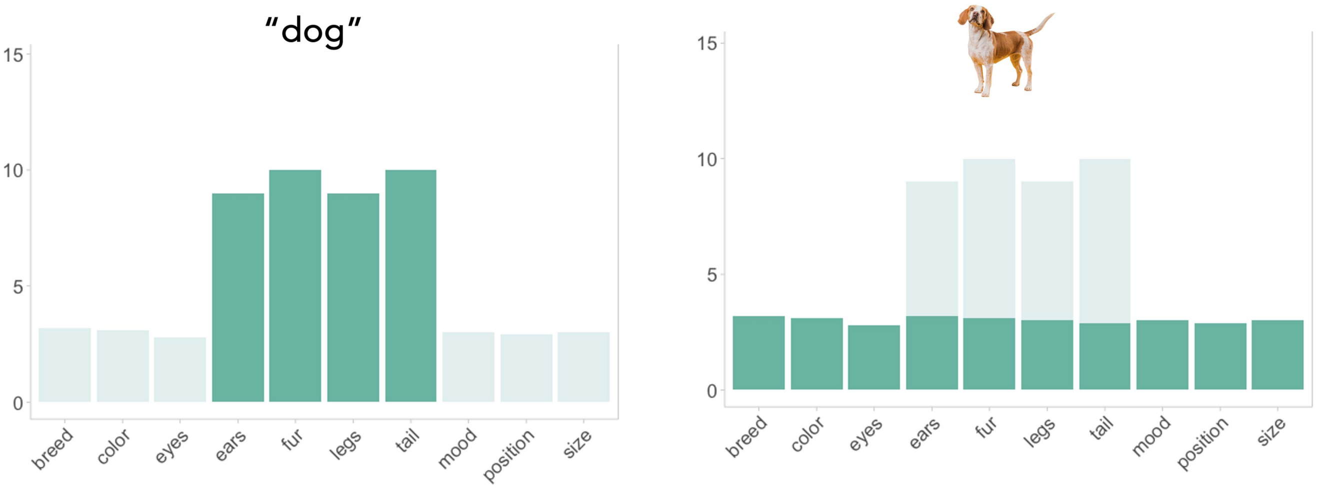
While these two lines of research—linguistic ability to spontaneously engage perceptual regions and linguistic (cueing) top-down influences on our perceptual experience—may not seem closely connected in the literature, they converge on the mechanism by which language engages the perceptual system. The peculiar ability of language to affect perceptual processes has been formalized as the label-feedback hypothesis by Lupyan (2012). This account proposes that linguistic labels (i.e., words) are the perfect vehicle for categorization. By being "unmotivated"—untied to a particular exemplar—linguistic labels create a tool for conceptualization that allows one to extract the diagnostic features of the lexicalized concept and the category it belongs to without the superfluous details (Edmiston & Lupyan, 2015). For example, while a picture of a dog or the pitch of a dog's bark will inevitably contain some information about the specific instance of the dog portrayed in the picture or through the sound of the bark (e.g., the size of the dog), the word "dog" will not automatically contain such information. For some more conventional categories, we might have found a way to systematically subcategorize certain concepts, like a dog breed, but even these subcategories are still untied to any particular instance of that breed, and are devoid of details tied to such instance (e.g., the color pattern of the fur), whereas that information is inevitably contained in a typical picture. That lack of nondiagnostic (i.e., less relevant for the definition of the category) details makes the category-diagnostic (i.e., relevant for category construal) information particularly prominent and allows for the creation of a strong, stable, and noise-free conceptual representation capable of engaging the perceptual system and biasing ("warping") its activation toward the labeled concept (for a detailed account of the label-feedback hypothesis, see Lupyan, 2012; Lupyan et al., 2020).

The two main claims of the label-feedback hypothesis—that language modifies perception in a feature-based manner and that the key to its efficiency lies in the absence of evoking representations with unnecessary details—are not new in the field of neuroscience, although they are rarely referenced in neurolinguistic research. The feature-based nature of top-down influences on perception has been well documented in studies on the effect of attention and expectation on perception (Maunsell & Treue, 2006; Saenz et al., 2002; Summerfield & de Lange, 2014; Summerfield & Egner, 2016). Namely, scientists have found that the encoding of components or features of visual input can be enhanced by means of top-down attentional and expectational influences. Building upon feature-based accounts of attention and expectation, the label-feedback hypothesis simply states that linguistic labels optimize the extraction of relevant (i.e., category-diagnostic) features (by not being tied to a specific exemplar), and in doing so end up being particularly efficient in providing feature-based top-down influence on perception. The detrimental effects of too many details, dubbed the dilution effect, have also been well documented in studies on decision making (Hotaling et al., 2015; Nisbett et al., 1981). The dilution effect (see Figure 1) describes how we weigh (the importance of) pieces of information in order to reach a certain decision or confirm a hypothesis. According to this account, when considering how much influence every piece of information has in informing our decisions, the claim is that our brains do not sum up, but rather average across pieces of information we receive. This means that the brain does not add up pieces of information according to "the more, the merrier" principle, such that more information would reinforce our conviction in our hypothesis. Rather, the more pieces of information the brain receives, the less weight each individual piece of information has, thus diluting the information crucial for category construal. In other words, superfluous details, even when accurate, dilute the weight of category-diagnostic features and end up creating a less strong conceptual (and perceptual) representation in the brain (Hotaling et al., 2015; Nisbett et al., 1981). By strength of representation of a stimulus in the brain, we refer to the extent to which a stimulus can engage neural networks commonly activated during the processing of said stimulus, and possibly exert an influence on other, subsequent cognitive and neural processes. As such, the strength of stimulus representation in a behavioral task might be reflected in the extent to which the encoding of the given stimulus can affect the processing of subsequent stimuli (e.g., speed up or slow down reaction times (RTs) when detecting subsequent stimuli, bias people toward seeing the subsequent stimuli [in]correctly). Neurally speaking, the strength of stimulus representation can be shown as an increase in neural activity in the regions in the brain responsible for the processing of that stimulus, as measured by various neuroimaging methods (e.g., higher levels of blood oxygen level dependent signal in the regions responsible for stimulus encoding or more widespread neural activity in the brain, larger event-related potential [ERP] amplitudes, or increased power of a frequency band associated with the processing of the stimulus).

From a neural perspective, this theoretical account would suggest that, as language enters the brain through language-specific sensory cortices (e.g., visual or auditory), the linguistic input is semantically/conceptually encoded alongside its diagnostic, categorical features in high-level semantic–conceptual areas such as the middle temporal gyrus (MTG), anterior temporal lobe (ATL), and inferior frontal gyrus (IFG). Shortly after, the visual features associated with the linguistically delivered concepts are activated in object- or feature-specific sensory regions to enrich the understanding of the said concept or further contextualize its implication. This neural process can then have further consequences for how we encode the incoming visual information, given that certain visual regions may already be preactivated at the time of the arrival of the visual input. This preactivation or "head start" of the neural activity can then have a facilitatory effect when encoding the upcoming input if that input is conceptually congruent with the ongoing process, or detrimental

Figure 1

Dilution Effect in the Context of Language–Perception Interaction



Note. People make predictions about the upcoming stimulus based on the information they have in the moments preceding that stimulus. When having to weigh several pieces of information (like several features or attributes of a concept), with different levels of relevance to the judgment at hand, the presence of irrelevant information dilutes or weakens the importance we give to the relevant information. For example, the word "dog" leaves many details unspecified and therefore makes the core features of a dog, such as the fact that they are four-legged animals with fur, a tail, and characteristic ears, quite prominent. A picture, on the other hand, inevitably provides details such as the color pattern, size, breed, and position of the dog, thus diluting the core features or characteristics of the concept "dog," making them less strongly activated in the brain. See the online article for the color version of this figure.

effects if that input is conceptually incongruent with the ongoing neural process (see Figure 2).

In sum, the current theoretical stance on linguistic effects on perception is that by being untied to any particular category exemplar—and therefore optimally objective and restricted to defining categorical features only—linguistic labels can engage perceptual systems in a particularly strong and noise-free manner, resulting in a stronger effect on any potential visual input. The basis of this theoretical postulation is in line with the (nonlinguistic) top-down and information-processing mechanisms reported in the cognitive neuroscience literature, which is an important step toward studying language in the context of general cognitive and neural mechanisms in the brain.

Review Overview

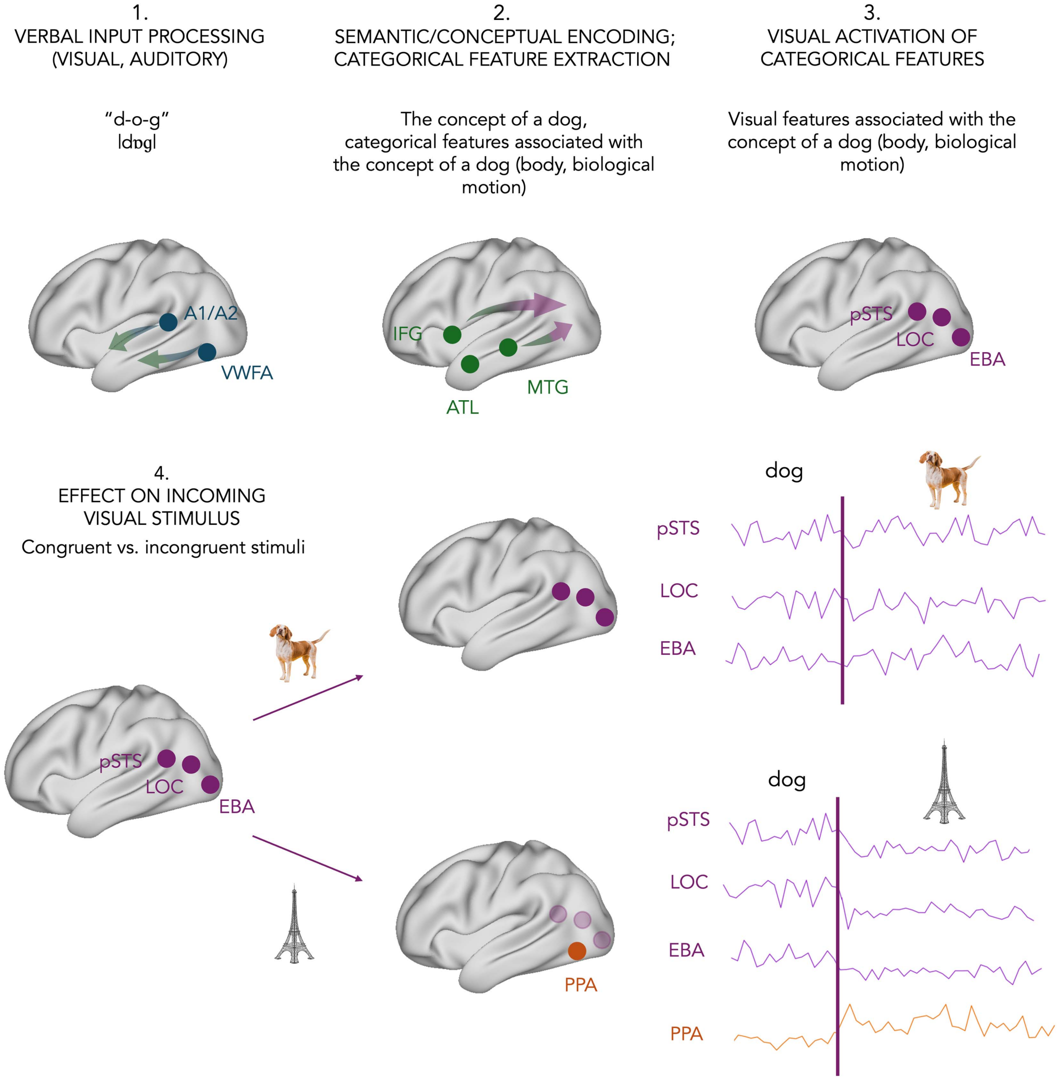
We have come a long way in understanding how language affects perception. However, the conditions necessary for linguistic effects on perception to be observed, as well as the exact neural mechanism behind this type of effect, are still under debate (Firestone & Scholl, 2016; Pylyshyn, 1999). This is the case for two reasons. First, lexical influence on perception can adopt many seemingly conflicting forms, depending on the choice and timing of stimuli, the task, or demands on the working memory, that might conflate any linguistically driven top-down effect. Second, while there are many behavioral studies showing both facilitating and interfering effects of language on perceptual judgment, studies looking into neural correlates of those behavioral effects are still scarce. However, with more research providing a nuanced profile of linguistic top-down influence on a wide range of perceptual tasks, using both behavioral and neuroimaging methods, a clearer picture is emerging about the nature of such effects.

This review aims to provide a summary of findings on the topic of linguistic influence on visual perception as reported in the literature, in light of the neural mechanisms underlying such processes. In order to do so, we will bring together literature from research investigating the scope of semantic representations in the brain and from research on linguistic (cueing) top-down influences on perceptual experiences. We believe that both of these approaches complement each other in constructing a comprehensive narrative on how language affects the perceptual systems. We therefore address the contributions from both lines of research by putting forward the notion that any observed shift in how we receive and interpret visual input as a function of language is a by-product of linguistic ability to evoke neural activity in perceptual regions. Furthermore, we will relate these findings to more generally observed mechanisms underlying nonlinguistic top-down (expectation and attention) influences in perception.

To that end, we will review studies looking at linguistic engagement of perceptual systems in the absence of a visual task or target, as well as studies using linguistic cues to modify the processing of the visual task or target. We will specifically focus on studies of visual perception, rather than on other types of perception (de Araujo et al., 2005; Herz & von Clef, 2001; Manescu et al., 2014; Okamoto et al., 2009; Orgs et al., 2006; Sohoglu et al., 2012; Sorokowska et al., 2015) or sensorimotor cortices (for the interaction between language and sensorimotor systems, see Hauk & Tschentscher, 2013; Pulvermüller, 2005; Pulvermüller & Fadiga, 2010). Seeing as we are interested in the neural mechanisms of linguistically mediated perceptual activation, we will only look at studies that focus on the immediate, short-term, in-the-moment effect of language on perception, rather than potential long-term effects of language (as in, e.g., cross-linguistic studies, comparing speakers of languages with different semantic categories, see Wolff & Holmes, 2011). However, we find it important to briefly acknowledge that different languages can guide the speakers to conceptualize the world around them slightly differently, and

Figure 2

A Schematic Overview of Neural Process Underlying Linguistic Influence on Perception



Note. Step 1: The linguistic information enters the brain via the sensory system (auditory: spoken language; visual: written language) and is received by and initially processed in the auditory cortex (A1/A2) or the visual word form area (VWFA). Step 2: The word meaning is then processed in the language–conceptual network. Since we give an example of a single word (i.e., linguistic label), the word is semantically processed in the left middle temporal gyrus (MTG), anterior temporal lobe (ATL), and/or inferior frontal gyrus (IFG). The processing of the word entails the encoding of the word's semantics and the concept it represents, as well as taking stock of and extracting diagnostic features associated with the concept at hand. Step 3: Categorical features of the concept are visually encoded shortly after in category/feature-selective regions within the visual cortex, such as the posterior superior temporal sulcus (pSTS) involved in biological motion perception, as well as lateral occipital complex (LOC) and extrastriate body area (EBA) involved in body perception. Step 4: If that process is then intercepted or followed by the arrival of a visual stimulus, the effect of the word on the visual stimulus is twofold. If the word and visual input are conceptually congruent (word "dog" followed by an image of a dog), then the word-evoked coactivation of visual features will facilitate the processing of the visual stimulus. If they are conceptually incongruent (word "dog" followed by an image of a landmark), then the ongoing word-evoked coactivation of visual features will need to be inhibited in favor of encoding the incoming visual stimulus (in this case, in the parahippocampal place area [PPA], responsible for the encoding of scenes and places). This process of course correction can lead to a delayed recognition of the visual stimulus or even false identification of the stimulus when the observer is under time pressure or is looking at a visually degraded or ambiguous stimulus. N.B. While this is a schematic view of bottom-up and top-down processes in the brain, it is important to acknowledge that feedforward and feedback connections do not just run between perceptual, linguistic, and conceptual regions, but are abundant across different levels of perceptual cortices, as well as across different language and conceptual regions. Likewise, the functional regions included in this schematic overview are simplified to illustrate the processing stages of language–perception information flow. Vs. = versus. See the online article for the color version of this figure.

therefore, which features might be diagnostic of a linguistically encoded concept may indeed be different across languages (Kemmerer, 2023; Slivac & Flecken, 2023). However, while different languages can vary in how concepts are being grouped into different categories, we argue that the cognitive and neural mechanisms stay the same across languages. In other words, speakers of different languages still go through the process of encoding a verbal stimulus, extracting diagnostic information from its semantics, and grounding or situating the relevant aspects (which may be dictated by the language itself, culture, task at hand, etc.) of the verbally expressed concept in sensory cortices, which can then facilitate the processing of visual information in the case of word-visual congruence, or interfere with the processing of visual information in the case of word-visual incongruence.

Semantic Representations in the Brain

Researchers interested in the scope of semantic representations in the brain have focused on studying the ability of language to spontaneously engage the visual system, while participants passively read or listen to stories, sentences, or words (Huth et al., 2016; Mathôt et al., 2017, 2019; Nijhof & Willems, 2015; Saygin et al., 2010; Wallentin et al., 2011). These studies usually involve an eye-tracking or neuroimaging method, without a behavioral task, given that they do not set out to measure the effect of language on behavior. The advantage of these studies is that the findings are not conflated with demands associated with the task. The disadvantage is that the researchers have little control over how much attention participants are paying to the stimuli during the experiment.

Using functional magnetic resonance imaging (fMRI), it has been shown that listening to stories evokes neural activation patterns encompassing cortical surfaces far beyond conventional language comprehension and semantic integration areas, such as the left IFG or the MTG and the left ATL (Huth et al., 2016; Nijhof & Willems, 2015). Not only did these studies show that linguistic inputs engage broadly distributed networks, but they have also highlighted the fact that the neural representation of any given word is both modality-specific (i.e., different clusters can be observed for words depicting actions, tools, colors) and influenced by the company the words keep in any given context. However, there is still more to be examined about the neural mechanisms underlying linguistic ability to engage nonlinguistic regions (for an overview of the short-comings of these types of studies, see Barsalou, 2017). Additionally, some of these studies might have capitalized on the evocative power of rich linguistic contexts—stories—to recruit featural information from perceptual regions, raising the question of whether contextually poorer stimuli, like isolated sentences or words, could still engage visual cortices in a similar manner.

To that end, more controlled studies have used carefully constructed sentences in order to examine whether their comprehension engages perceptual processes (Saygin et al., 2010; Wallentin et al., 2008). Saygin et al. (2010) used fMRI to investigate whether motion sentences can modulate visual motion perception regions in the brain, such as the middle temporal area (MT/V5), during natural language comprehension. They found a gradual shift in the activation of the MT/V5 region, with motion sentences exerting the strongest influence on the region, fictive motion sentences showing a slightly lesser effect, and static sentences having no effect on the MT/V5 activation. These results show a nuanced profile of the ability of language to modulate visual regions, with even metaphorical concepts having an effect on the visual system by means of their literal meaning and the corresponding perceptually encoded features, albeit less strongly than literal sentences.

Looking at the effects of isolated words on the visual system in the absence of visual stimuli, studies have shown that even linguistic labels in isolation can engage perceptual cortices, such that the features encoded in the semantics of the labels activate regions close to or overlapping with the cortical areas that mediate the perception of those features (Martin et al., 1995; Rueschemeyer et al., 2010; Simmons et al., 2007). Mathôt et al. (2017: original study in French, 2019: replication study in Dutch) measured participants' pupillary dilation and constriction in response to auditorily presented words conveying the notion of darkness (e.g., night, shadow) or light (e.g., day, sun). The authors observed pupillary dilation in response to darkness-conveying words, and constriction in response to light-conveying words, suggesting that a word's semantics modulates the activation of cortical areas responsible for pupil control in an involuntary manner (Mathôt et al., 2017, 2019).

Interpreting Results: Semantic Representations in the Brain

Taken together, these studies show us that when reading or listening to linguistic stimuli, language spontaneously engages functional networks that surpass conventional language regions in temporal and frontal cortices. Specifically, they have demonstrated that linguistic stimuli, from elaborate stories to words in isolation, can recruit perceptual cortices and activate regions that usually respond to visual stimuli (but in this case, they respond to the visual representations of the features encoded in the word semantics).

While these studies show that language can evoke neural activation in perceptual regions, even in the absence of a visual stimulus, they cannot tell us what the behavioral and neural consequences of such modulations are in the presence of a visual input. In other words, they cannot tell us how these modulations change our perceptual experiences and the way in which we encode incoming visual input. In order to tackle that question, we next look at the studies examining whether and to what extent language can modify how we receive and interpret visual input and how those modifications are actualized in terms of neural activation patterns.

Cueing Effects of Language on Perception

Studies interested in the ability of language to modify ongoing perception and perceptual judgments usually utilize a cueing paradigm. In this paradigm, people are presented with a linguistic stimulus, such as a word, sentence, or short story, prior to doing a perceptual task, such as detection, discrimination, or a visual search task (i.e., identifying the target among multiple distractors) on the target (Boutonnet & Lupyan, 2015; Dils & Boroditsky, 2010; Francken, Kok, et al., 2015; Meteyard et al., 2007; Slivac et al., 2021). By intercepting the process of language comprehension with a visual input, these studies can directly examine how the linguistic engagement of the visual system, observed in studies with purely linguistic inputs covered above, affects the processing of visual information, both behaviorally and neurally. The benefit of these studies is that they, by virtue of having an explicit task, have more control over what participants pay attention to. Additionally, by connecting behavioral effects with the neural pattern of activation in response to experimental conditions, they provide a more comprehensive account of language–perception interaction and tap into possible mechanisms that lead to those interactions. The disadvantage of such experiments is that the task may conflate the effects of language on perception by biasing the attention or expectation of participants to certain aspects or features of a concept (Hoenig et al., 2008; Kerzel et al., 2009; Yee et al., 2012; Yee & Thompson-Schill, 2016).

Behavioral Studies

Studies investigating how language comprehension affects our visual processing have reported both facilitating effects on target perception, when language cues are congruent with the target, and detrimental effects in the case of cue–target incongruence, suggesting that language has the power to change our perceptual experiences, that is, modify how we receive and interpret visual input.

Studying the effect of motion stories on visual motion perception, Dils and Boroditsky (2010) found that listening to linguistic descriptions of motion could induce mental images strong enough to cause a motion aftereffect illusion—an illusion of a directional motion in a static or incoherent random dot motion stimulus caused by prior exposure to motion in the opposite direction (Dils & Boroditsky, 2010). This finding demonstrated that linguistic processing could change how we perceive subsequent visual input to such an extent that it can induce an illusion usually only triggered by visual stimuli. These studies again profited from rich linguistic input, suggesting that the potential visual shift could have been the product of participants' imagery in response to language abundantly reinforcing a categorical feature (e.g., upward motion).

Studies using more concise linguistic stimuli have also investigated the effect a single sentence can have on the perception of visual stimuli. These studies observed faster and/or more accurate responses to visual stimuli matching the content of the previously read sentence, such as motion direction, object orientation, and shape (Pelekanos & Moutoussis, 2011; Stanfield & Zwaan, 2001; Zwaan et al., 2002).

Using still more concise lexical stimuli, studies using single words (i.e., linguistic labels) have observed some diverging results. Using linguistic labels as spatial cues, some studies have confirmed the label advantage in the form of faster reaction times for spatially congruent target detection (Dudschig et al., 2012; Taylor et al., 2015; Weger & Pratt, 2008) and discrimination (Chasteen et al., 2010; Ostarek & Vigliocco, 2017). Conversely, other studies have shown that language spatially congruent with the target can have a detrimental effect in the form of slower responses to the target (Estes et al., 2008; Richardson et al., 2003; Verges & Duffy, 2009). Such detrimental effects have been explained by arguing that the process of identifying a target in the cue–congruent location may first require an inhibition of spatial features activated by the cue (Estes et al., 2015). Similarly, it has also been hypothesized that the inhibition effect in those studies is due to the feature overlap being manipulated between the cue and the location of the target, rather than the cue and the target itself (Dunn et al., 2014), or due to the fact that the cueing effect interfered with the more demanding discrimination rather than a simpler detection task (Dudschig et al., 2012). However, given that the two studies that reported the cue-induced spatial facilitation effect used spatially congruent but cue-irrelevant targets (Dudschig et al., 2012) and a discrimination task (Ostarek & Vigliocco, 2017), more research is needed in order to understand the conditions and mechanism behind the inhibition effects of cues spatially congruent with the target location. Finally, it has also been suggested that timing, alongside the nature of the task (detection vs. discrimination), might play a crucial role in observing spatial inhibition or facilitation effects, with shorter stimulus onset asynchronies (SOA) leading to inhibition and longer SOA leading to facilitation of target perception (Gozli et al., 2013). Building upon Chasteen et al. (2010) and Estes et al. (2008) studies, Gozli et al. (2013) found that, in the discrimination task, a short SOA leads to an interference effect, whereas a long SOA leads to a facilitation effect. Additionally, they found that the presence of single or multiple conceptual categories in the cueing studies plays a role in the type of effect to be observed, such that in a single category experiment, they always (for short and long SOA) observed a facilitation effect, whereas in the multiple categories experiment, they observed the early (short SOA) interference versus late (long SOA) facilitation dichotomy. Taken together, these studies reveal the complexity of the role of different factors and their interactions that might play a role in observing inhibitory or facilitatory influences in linguistic effects on perception.

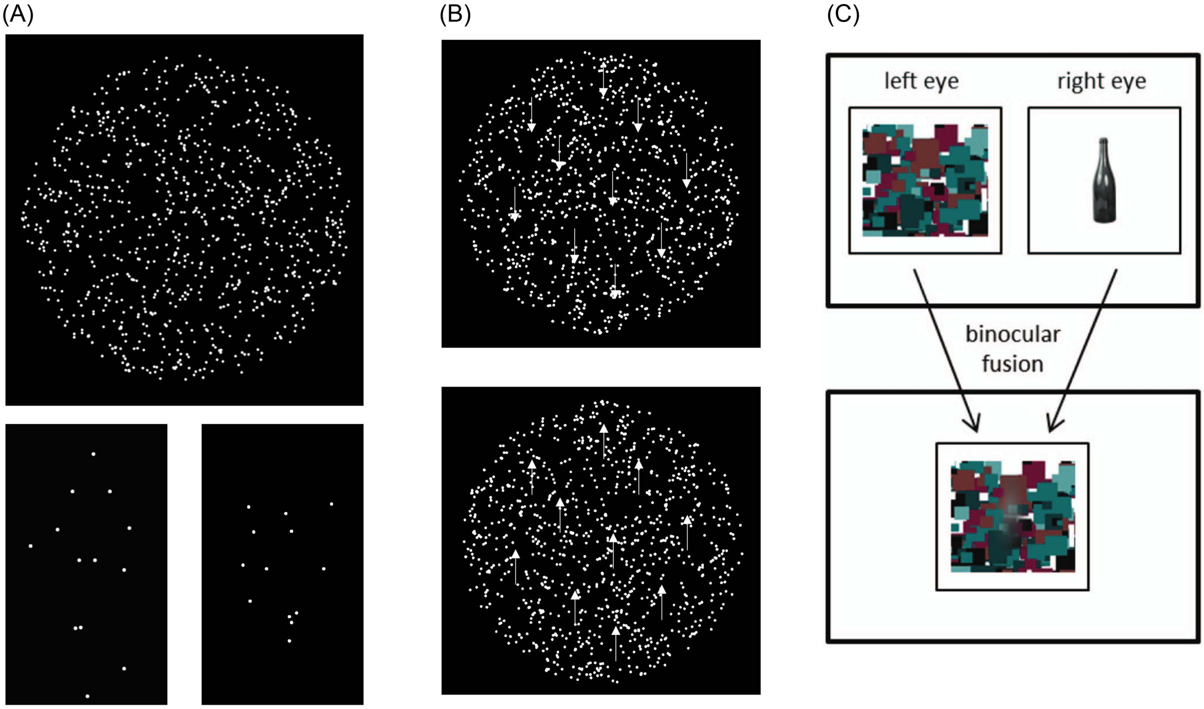
Moving beyond spatial cues, numerous studies have demonstrated that the presentation of a single word (i.e., lexical cue) shortly before a visual target can improve (in the case of cue–target congruence) or interfere with (in the case of cue–target incongruence) target discrimination (Forder & Lupyan, 2019; Lupyan & Thompson-Schill, 2012; Ostarek & Vigliocco, 2017), detection (Bidet-Ildei et al., 2011; Edmiston & Lupyan, 2015; Francken, Kok, et al., 2015; Meteyard et al., 2007; Pavan et al., 2013), or visual search (Lupyan & Spivey, 2010b; for review, see Huettig et al., 2011).

This effect has been shown to persist even when the visual target is rendered imperceptible or ambiguous by means of masking techniques such as continuous flash suppression (Forder et al., 2016; Lupyan & Ward, 2013; Ostarek & Huettig, 2017; Pinto et al., 2015; Stein & Peelen, 2015), random dot motion masking (Bidet-Ildei et al., 2011; Slivac et al., 2021), or thresholding of motion coherence levels (see Figure 3; Francken, Kok, et al., 2015; Meteyard et al., 2007; Pavan et al., 2013). The design of these studies allows researchers to examine the extent to which we depend on higher level systems to fill in the gaps in our perception, and the results emphasize the reliance on the top-down linguistic information for successful target perception in visually noisy environments.

Interestingly, several studies have found the linguistic effect on perception to be left-lateralized, that is, observed when stimuli were presented to the right visual field, but not the left (Francken, Kok, et al., 2015; Zhou et al., 2010). Given that language processing itself is left-lateralized, these findings suggest that the strength of linguistic effects on perception may be dependent on the (hemispheric) proximity of the perceptual region to the linguistic areas. So far, these results are in line with studies reporting that long-term effects of categorization as learned through language on perception are also left-lateralized (Drivonikou et al., 2007; A. L. Gilbert et al., 2006; Mo et al., 2011; Regier & Kay, 2009). It has also been argued that the left-lateralized signature of categorical perception effects may not be driven by language, but rather can reflect a more general, nonlinguistic propensity of the left hemisphere for categorical processing. Specifically, Holmes and Wolff (2012) showed that categorical processing even for novel (and therefore unlabeled)

Figure 3

Types of Masking Techniques Used in Cueing Studies Investigating Lexical Cueing Effects on Perception



Note. Cueing studies usually involve a presentation of a verbal cue, quickly followed by a visual target upon which a decision needs to be made. Occasionally, studies will opt for a masked or visually degraded target to maximize target ambiguity and the effect the cues may have on target perception (usually either detection or discrimination). (A) Motion perception (biological): point-light figures of a human being engaged in the action of walking (coherent figure—left image, scrambled figure—right image) hidden in an aperture of randomly moving dots. In this type of study, participants have been asked to indicate whether the figure hidden in the middle of the aperture depicts a coherent human body engaged in a certain action (here, walking) or not. (B) Motion perception (direction): Random dot motion stimuli with a certain percentage of dots moving in one direction, while the rest of the dots move randomly. Participants are then asked to identify the dominant vector of the moving dots. (C) Object perception: Continuous flash suppression masking a 2D image of an object, where the target image is presented to one eye and the continuous flash suppression mask to the other eye. Participants are then presented with an auditory single-word stimulus and asked to indicate whether they see a specific object or not. See the online article for the color version of this figure.

object categories was more pronounced in the left hemisphere than in the right hemisphere, suggesting that the left-lateralized effect is equally induced by linguistic and nonlinguistic factors. Conversely, however, Witzel and Gegenfurtner (2011) found no lateralization when it comes to the effects of categorization on color perception, leaving the issue regarding the lateralization of category advantage in perceptual processing and its relationship to language up to further research.

Further, studies have shown that even cues unrelated to the task (Lupyan & Thompson-Schill, 2012; Slivac et al., 2021) or rendered unaware by means of masking (Francken, Meijs, et al., 2015) can modify how we process visual input. Additionally, Slivac et al. (2021) have shown that even cues unrelated to the task or the target, but rather congruent with the distractor stimulus, can still be detrimental to the detection of the target, suggesting that the distractor cues were strong enough to induce a bias away from the target. However, the cueing effects observed with subconscious or task-irrelevant cues are less pronounced than those with task-related or overtly shown cues, suggesting that the clear cue–task relatedness can reinforce the extraction of the features encoded in the cue semantics.

However, studies have also pointed out that contextual and task demands can substantially change the type of information extracted from word meaning (Hoenig et al., 2008; Yee et al., 2012; Yee & Thompson-Schill, 2016). Taken together, the evidence suggests that a task closely connected to the cue semantics can alleviate the computational burden by prioritizing the processing of that subset of lexical information deemed to be of the highest relevance for the task. While lexical influence on perception can survive task deviations from the semantic information encoded in the cue, it is still unclear to what extent shifting attention away from the target brings about a more substantial change in the kind of features that end up being extracted from the linguistic labels.

In sum, current behavioral evidence shows that language can shift how we receive and interpret visual inputs. Namely, hearing or reading linguistic cues featurally congruent with a visual target facilitates the perception of that target, and it can even boost the target detectability threshold (i.e., make the invisible targets visible). Conversely, in the case of cue–target incongruence (i.e., lack of shared features), language has a detrimental effect on our ability to process visual information quickly and accurately. Current evidence also shows that the choice of task can up- or downregulate the potency of lexical effects on perception by highlighting the relevance of the semantic content conveyed by the cue needed for the successful performance of the task.

Interpreting Results: Behavioral Studies

Linguistic influence on perception can manifest itself in a number of different ways. Cueing studies using a visual detection or discrimination task investigating this phenomenon have interpreted findings pertinent to accuracy scores, reaction times, and signal detection theory indices (discriminability/sensitivity and bias) as proof of linguistically mediated facilitation of perceptual processing.

To illustrate, comparing cue–target congruent with incongruent experimental conditions, motion perception cueing studies found a facilitating effect of congruent motion cues on motion perception, as reflected in both an increase in accuracy and faster reaction times (Francken, Kok, et al., 2015; Slivac et al., 2021), faster or slower reaction times only (Bidet-Ildei et al., 2011; Richardson et al., 2003), a shift in both d' and criterion (Meteyard et al., 2007), and a shift in criterion only (Francken, Kok, et al., 2015; Slivac et al., 2021). Although these findings may seem somewhat inconsistent, the explanation for this inconsistency may lie in the experimental settings themselves.

One possible explanation for some of these discrepancies in findings comes from the study done by Pavan et al. (2013), who tested if and how lexical effect on perception changes as a function of an increase in target ambiguity. Presenting participants with random dot motion targets with coherence levels at suprathreshold (84%) and threshold (50%) accuracy levels, they found a double dissociation between discriminability (d') and reaction times for the two types of targets. For the suprathreshold condition, the reaction times were faster for directional cue–target congruence, but there was no difference in discriminability. For the threshold condition, no reaction time effects were found for cue–target congruence, but there was a shift in d' (Pavan et al., 2013). The theoretical explanation for this dual dissociation may be that highly detectable stimuli do not necessitate reliance on top-down influences for a successful target detection. In other words, priors may have the strongest influence on perception when visual input is ambiguous (Bogacz et al., 2006). Additionally, given the high accuracy performance on tasks with highly visible targets, there may not be much space to significantly increase accuracy as a function of facilitating cues compared to the baseline condition, especially given attention lapses and time constraints regularly present in cueing experiments. In those cases, the facilitation might manifest itself in reaction times instead. However, in the case of ambiguous stimuli, our reliance on top-down information for target perception becomes much more meaningful. In addition, there is a lot more space to improve performance on the task as a function of congruent cues. However, the time necessary to resolve ambiguity might not be significantly changeable within the restricted time window that cueing detection studies usually have. More comprehensive measures such as speed–accuracy trade-off can potentially provide further insight into how the brain processes visual information as a function of congruent and incongruent lexical cues.

Additionally, signal detection theory indices, such as sensitivity/discriminability (most often expressed as d') and bias (most often expressed as criterion) measures are often reported as being indicative of the cognitive and neural loci of the linguistic top-down influence on perception. Sensitivity measures tend to be interpreted as a sign of perceptual modulations, while bias has been taken to be an indicator of higher level, decision-making processes (Meteyard et al., 2007; Pelekanos & Moutoussis, 2011). This dichotomy, however, has been brought into question by a number of studies, which pointed out that signal detection indices must be interpreted in the context of experimental settings, such as stimuli and tasks (Georgeson, 2012; Witt et al., 2015). Furthermore, it has been shown that shifts in bias are accompanied by modification in activation patterns in perceptual regions. Several studies have presented convincing results showing that criterion in discrimination tasks is either sensory in nature, or both sensory and decisional, but not solely decisional (Linares et al., 2019; Morgan et al., 1990; Ratcliff et al., 1989).

Even knowing that both sensitivity and bias shifts stem from activity in visual cortices, the discriminability-bias dichotomy does have an important role—it can tell us much about the behavioral consequences of lexically induced perceptual shifts (see Figure 4). One possibility, reflected in changes in discriminability, would be that linguistic activation of visual regions helps us differentiate between different stimuli, that is, makes us more adept at distinguishing between the target and nontarget. Another possibility, reflected in bias shifts, is that language warps visual perception toward the cued concept, such that under the influence of a target–congruent cue, we are less successful at discriminating between two stimuli and more likely to classify a nontarget as a target, especially in ambiguous settings.

Summerfield and Egner (2016) proposed that the distinction between discriminability and bias can be connected to the feature-based attentional and expectational top-down influences. They claim that feature-based attentional top-down processes manipulate the relevance of the incoming visual signal, such that they upweight the information relevant for the task during the decision-making process, resulting in discriminability (i.e., sensitivity) shifts. Feature-based expectation, on the other hand, affects the probability of the incoming signal, thus adjusting the decision criterion toward the more probable features, that is, those that occur more frequently in any given context (Summerfield & Egner, 2016). Even in nonlinguistic studies on top-down influences on perception, these two aspects have rarely been explicitly orthogonalized. In studies involving language, this is particularly difficult to achieve, given that language semantics can be a complex and multifaceted source of information for conceptualization and feature extraction.

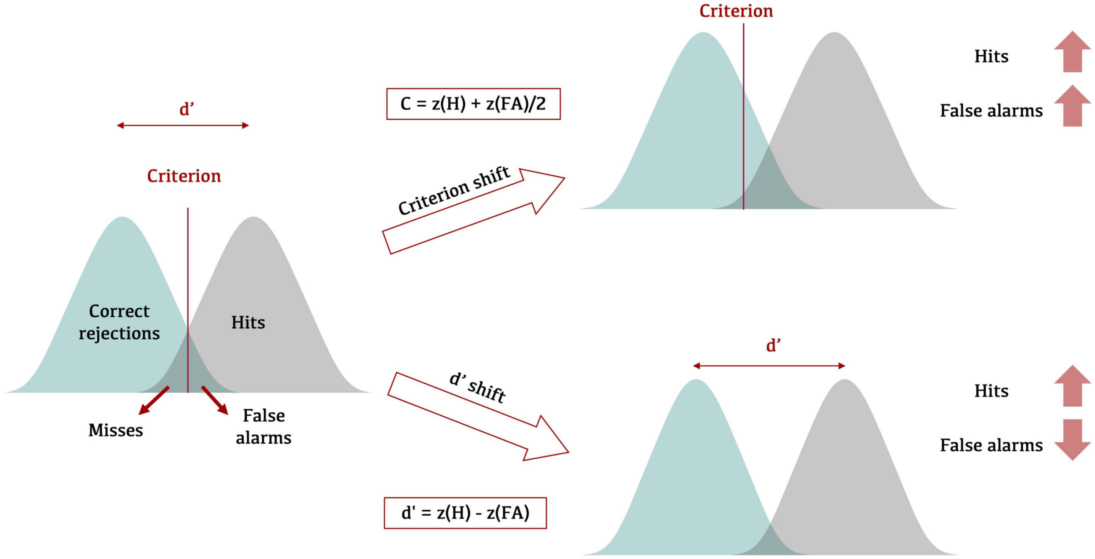
Neuroscientific Studies

While studies looking at the linguistic ability to spontaneously engage perceptual systems have relied mostly on neuroscientific techniques in their experiments, cueing studies with perceptual tasks, however, have mostly employed behavioral paradigms, with very few attempts to study neural activation patterns accompanying the behavior. Still, the presence of eye-tracking and neuroimaging techniques is not completely absent, so we will take a look at the current state of knowledge about the neural correlates of linguistic effects on perception.

These studies tackle the question of whether linguistic top-down processes change the neural representation of visual stimuli in perceptual areas, that is, whether the integration of lexical information and bottom-up visual inputs can already be observed at perceptual

Figure 4

Two Possible Response Profiles of Linguistic Influence on Perception



Note. *Criterion shift:* In a target detection task, people are more likely to report seeing the target when the cue is congruent with that target. However, this bias might be part of a more general tendency to report seeing whatever is encoded in the cue, regardless of the target. While we might observe an increase in the hit rate when the cue is congruent with the target, at the same time, we might also observe an increase in the false alarm rate when the cue is incongruent with the target, indicating a strong tendency to see what has been cued, which comes at the expense of accurately perceiving what has been visually presented. *d′ shift:* People are more adept at discriminating between the target and the nontarget when cued by the target–congruent cues—they are better at detecting the target, but also better at correctly rejecting the nontarget. So the cue might function as an amplifier of relevant features for a given concept and aid in accurately perceiving both the presence and the absence of those features in the visual stimuli. H = hit; FA = false alarm. See the online article for the color version of this figure.

processing stages. Alternatively, perceptual processes may be unaffected by top-down lexical influences, and the integration of lexical information and perceptual processing may be happening at the higher, decision-making stage instead.

Eye-tracking studies using words conveying spatial information have shown that linguistic cues, implicitly encoding the spatial feature congruent with the target location, can facilitate saccadic eye movements (initiation and speed) toward that location (Dudschig et al., 2012, 2013; Dunn et al., 2014). These results suggest that words may reach neural processes responsible for the control of eye movements and pupillary responses.

Electroencephalography (EEG) measurements, given high temporal resolution, can tell us whether the label advantage stems from modulations in the early visual processing level or later, semantic and decision-making level, by looking at whether lexical cues elicit early visual ERP components, such as the P100 and N100, or a later ERP component, such as the N400, associated with higher level semantic integration (Boutonnet & Lupyan, 2015; Landau et al., 2010; Noorman et al., 2018). Landau et al. (2010) examined the temporal dynamics of the linguistic effect (sentences) on face perception and found that the magnitude of the N170, associated with face perception, was larger after face-describing sentences than after scene-describing sentences, and only for the left hemisphere. Looking at the top-down effects of linguistic labels on perception, Boutonnet and Lupyan (2015) found that hearing a word affected early visual processes, as reflected in the P100 ERP component associated with the processing of low-level visual features. Namely, they found that the P100 was larger when participants were cued by labels (e.g., the word “dog”) compared to equally informative nonverbal cues (e.g., dog bark), with the enhancement occurring within 100 ms of image onset. Similarly, Noorman et al. (2018) showed that hearing a word activates the shape representations of its referent, which affects the visual processing of a following picture within 100 ms from its onset. These findings are compatible with the general view that features extracted from cues act as immediate priors for the visual system, thus immediately biasing the reception and processing of the incoming visual target.

Using sentences with content matching or mismatching perceptual features of target objects in a magnetoencephalography (MEG) study, Hirschfeld et al. (2011) found two effects: an early modulation in the occipital cortex within 120 ms, reflecting early visual processing, and a later modulation in the N400 window in the left temporal cortex, sensitive to higher level processes such as lexical access and semantic integration. These results show that the loci of the linguistic top-down influence on perception can be detected at both lower perceptual levels as well as higher semantic levels.

Studies using fMRI have looked at the extent to which lexical cues can modify neural activation in perceptual regions responding to the visual target, as reflected in changes in the blood oxygen level dependent signal (Francken, Kok, et al., 2015; Pirog Revill et al., 2008; Puri et al., 2009). Puri et al. (2009) employed a cueing task in an fMRI setting to investigate whether linguistic labels (face vs. place) cue the perception of faces and places, known to be represented in the fusiform face area and parahippocampal place area, respectively. They found that cue–target congruence led to an increase in the activity in the target-encoding regions, compared to the incongruent pairings. These findings showed that linguistic labels can modify neural activity in cortical areas responsible for encoding visual features, which in turn suggests that language can cascade into the perceptual system and modify neural activity therein. Conversely, Eger et al. (2007) used degraded visual stimuli to examine the effects of lexical cues on object perception and found that even though objects congruent with the cue were recognized earlier than incongruent ones, when controlling for recognition point and stimulus information, activity in the ventral visual cortex reflected recognition success, independent of the cueing condition. Similarly, Francken, Kok, et al. (2015) looked at the cueing effects of motion words on the perception of motion directionality and found the locus of the linguistic effect on motion perception in the MTG, involved in semantic integration, rather than the hypothesized MT/V5 region in the extrastriate cortex, involved in motion perception. Taken together, fMRI research presents a divided picture, without clear evidence for whether linguistic influence on perception can be detected in earlier visual cortices using the fMRI technique.

Interpreting Results: Neuroscientific Studies

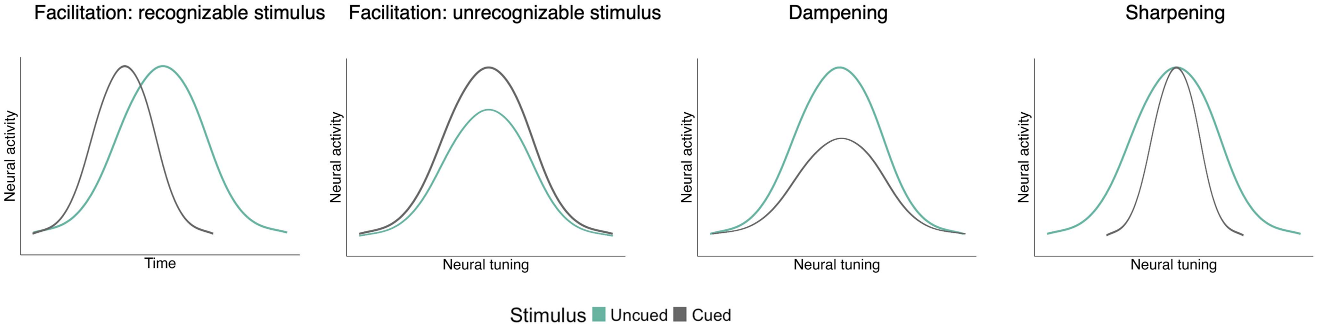
Neuroscientific cueing studies show divergent results between M/EEG and fMRI results. Namely, M/EEG studies are consistent in showing the early influence of linguistic cues on perception, while this observation is largely absent from the fMRI studies. This is particularly surprising given that studies looking at the representation of semantic maps in the brain show an engagement of perceptual regions corresponding to visual representations of categories and features depicted by language.

Given that linguistic top-down effects on perception happen at a very short timescale, it could be that the low temporal resolution of fMRI does not allow us to reliably capture these language-evoked, transient boosts in activation in sensory areas. While there is also a possibility that any visual input overpowers the neural activity in the perceptual cortices evoked by language alone, M/EEG and behavioral findings suggest that this is not a likely scenario. We should note that none of the fMRI cueing studies reported above had a control condition without a lexical cue (i.e., all cues, congruent, incongruent, and neutral, when present, were lexical). One possible scenario, therefore, is that lexical cues do shift neural activation in perceptual regions compared to nonlexical cues, but the difference in cue congruence between different lexical cues themselves is not reflected in an overall increase or decrease of activation within a visual region of interest, as measured with standard univariate fMRI analyses approaches. Further, the effect of lexical cues congruent with the target, compared to those incongruent with the target, could only be detectable by looking at the differences in more fine-grained activation patterns within perceptual regions of interest, detectable with multivariate pattern analysis approaches (i.e., looking at subtle shifts in neural responses that take into consideration the patterns of activation across voxels).

Several possible activation profiles have been proposed as viable signatures of (lexical) top-down influences on visual perception (see Figure 5). One possibility could be that the top-down influences evoked by lexical cues, when conceptually congruent with the visual target, will lead to an overall increase in activation in regions that encode that target—so-called facilitation or neural enhancement (C. D. Gilbert & Li, 2013; Müller et al., 2013). We suggest that neural enhancement could have two slightly different manifestations, depending on the recognizability of the visual target. If the target is known, unambiguous, or not visually degraded in any way, and already evokes maximal levels of activity that the target-selective neurons can reach, the peak activation of those neurons could be reached sooner, due to the head start already created by the cue. Facilitation accounts have suggested that these earlier peaks might also be accompanied by a shorter duration of neural activity in target-selective neural populations (Gotts et al., 2021; Grill-Spector et al., 2006; R. N. Henson, 2003; James et al., 2000). However, if the target is ambiguous, visually degraded by means of masking or subthreshold timing, or novel, the target–congruent cue can push the activation levels of target-selective neurons higher than the visually compromised target itself can (R. Henson et al., 2000; Müller et al., 2013; Turk-Browne et al., 2007). In this case, cueing would lead to an overall increase in target-selective neural populations, compared to an uncued target perception. Another possibility is that the top-down influence evoked by target–congruent cues can reduce the overall activity in regions encoding visual targets. This phenomenon is known as dampening, repetition or expectation suppression, attenuation, or neural fatigue, and proposes that the overall strength of activation of target-selective neural populations might be decreased as a function of expectation. According to this account, the most target-sensitive neurons end up being the most suppressed due to cueing. This leads to a decrease in activation in regions encoding the expected target when those expectations are met (i.e., in the case of the cue–target congruence in this context; de Gardelle et al., 2013; Grill-Spector et al., 2006; Müller et al., 2013; Summerfield et al., 2008). Finally, a more intricate version of the two previous accounts has been proposed: fMRI studies investigating top-down influences of nonlinguistic cues on perception have shown that target–congruent top-down cueing effects lead to the suppressed responses of the unexpected or target-irrelevant neural signal while simultaneously maintaining the strong neural response to expected or target-relevant features. This interplay between dampening of the unexpected or irrelevant neural signal and continued robustness of the expected or relevant neural signal results in the so-called sharpened pattern of activation across the target-selective neural population, rather than an overall increase or decrease in the average magnitude of activation in target-encoding regions (Desimone, 1996; Kok & de Lange, 2014; Kok et al., 2012; Lee & Mumford, 2003; Wiggs & Martin, 1998). In Bayesian terms, the attenuation, and by extension sharpening, accounts have been explained by the findings that the subpopulation of neurons coding for prediction error signals diminish, while those coding for expected stimuli (predictions) enhance, as a function of the expectation build-up, leading to an overall sharper representation of expected stimuli. The sharpening activation profile may not be observable with the standard, univariate fMRI analysis, which looks at the change in the average magnitude of activation within a region by evaluating each voxel in isolation, rather than patterns of activation within a region that take into consideration joint activation of multiple voxels. While the sharpening account of top-down influences on perception has been reported and argued for in a number of studies

Figure 5

A Schematic Illustration of Potential Neural Mechanisms Underlying Linguistic Influence on Perception (Cue–Target Congruence)



Note. Timing (when relevant) or the tuning curves of populations of sensory neurons in visual regions are shown in response to uncued (green) and cued (gray) visual stimuli (i.e., visual targets in the cueing paradigm). *Facilitation (recognizable stimulus):* If the visual target is clearly visible (unambiguous or unmasked) and familiar, there might not be a difference in the height of the peak activation between the cued and uncued target. However, a linguistic cue presented shortly before the target can already preactivate that same visual region (i.e., elicit activation in the visual region prior to target perception), making the peak activation of the target-selective neurons occur earlier than the target alone would. According to the facilitation account, the peak activation might not only occur earlier, but also last shorter. *Facilitation (unrecognizable stimulus):* If the stimulus is ambiguous, rendered invisible via masking, or unfamiliar, it might be too weak or unrecognizable to fully activate target-selective neurons in visual regions. In that case, the presentation of a target–congruent linguistic cue could serve to amplify the otherwise weak representation of the low-visibility or low-recognizability target. *Dampening:* Target-selective neurons might exhibit a reduced response to a cued target compared to an uncued target. According to the dampening account (also referred to as suppression or fatigue in the literature), cueing might reduce the overall response of target-selective neurons in the visual cortex, with the most target-sensitive neurons being the most suppressed due to cueing. *Sharpening:* According to the sharpening account, cueing may lead to a sparser representation of stimuli. In other words, cueing may narrow the tuning bandwidth of neural populations, such that those neurons that encode the critical information about the target maintain their robust activation levels, while those neurons encoding less relevant information about the target exhibit reduced activation. See the online article for the color version of this figure.

using nonlinguistic stimuli (Kok & de Lange, 2014; Kok et al., 2012; Martens & Gruber, 2012; Yon et al., 2018; Zhang & Sejnowski, 1999), it has yet to be tested in studies with linguistic labels as cues.

To summarize, M/EEG studies looking into language–perception interaction reliably show that linguistic cues can affect low-level visual systems as soon as 100 ms post target presentation. fMRI studies, however, are more divided when it comes to neural activation correlates of language–perception interaction. Not many studies have been done on the topic, however, and those that have been done report univariate results only, which can only measure the increase (or decrease) in the average magnitude of the signal in any given region, but tell us nothing about the more subtle shifts in activation within those regions. Nonlinguistic cueing studies, on the other hand, have shown that the effects of top-down expectations on perception may be accompanied by the sharpening rather than uniform increase or decrease of activation in early visual regions. This suggests that the fMRI studies may have been searching for the wrong or at least incomplete neural signature of the behaviorally observed linguistic cue–target interaction effect.

Discussion

In this review, we have provided an overview of the current state of knowledge regarding linguistic effects on perception, by summarizing findings from the literature investigating semantic representations in the brain and linguistic (cueing) top-down influences on perception. In the introduction, we have proposed a claim that linguistic top-down influences on perception observed in cueing studies are a by-product of a distributed semantic network, resulting in linguistically evoked neural activity in the perceptual cortices. In other words, cueing studies simply look at how linguistically evoked activity in perceptual regions modifies how we receive and interpret visual inputs.

The literature without a visual task has reliably shown that the processing of linguistic stimuli, from stories and sentences to individual words, can activate perceptual regions responsive to visual representations of categorical features conveyed by lexical items. These findings are in line with the mechanisms proposed from both linguistic and nonlinguistic top-down influences on perception: Perceptual system can be activated by cognitive (in this case linguistic) processes in a feature-based manner (Lupyan, 2012; Summerfield & Egner, 2016). In other words, part of the visual system responsible for the encoding of visual features representative of a category can also be engaged in the absence of visual inputs, during semantic processing of that category. But these findings beg the following question: How does spontaneous engagement of perceptual system affect how we then process subsequent visual information?

Behavioral cueing studies with visual targets have shown that this process of feature-based perceptual engagement can change how we receive and interpret visual inputs. The exact behavioral consequences of that influence, however, are not as consistent as one would think just looking at the literature examining whether linguistic processing spontaneously engages perceptual systems. In particular, studies investigating the effect of cue–target spatial congruence on target detection and discrimination have reported both facilitatory and inhibitory reaction times as a function of spatially congruent cues (Chasteen et al., 2010; Estes et al., 2008; Gozli et al., 2013, 2016; Ostarek & Vigliocco, 2017; Richardson et al., 2003; Taylor et al., 2015; Weger & Pratt, 2008). Various studies have tried to provide an explanation for the inhibitory results (see above), but the crucial reason for this discrepancy seems to be the extent to which the cue and target overlap in categorical features.

These findings illustrate what we propose in this review: In studies reporting inhibitory effects, the cue-evoked influence was in the location, not in the target itself, meaning that these studies overlapped two visual processes: the spatial influence put forward by the cue and visually reinforced by the target location, and the processing of the visual target itself, unrelated to the cue or the features characterizing the location. In this situation, the spatial encoding evoked by the cue would indeed need to be inhibited before the encoding of the target (i.e., an unrelated visual input) can happen, causing a delay in the processing of that target.

Studies using cues featurally overlapping with targets, on the other hand, present a consistent body of evidence showing that lexical cues can facilitate the perception of the visual target. However, the question remains about the neural signature of such facilitation. One possibility is that linguistic cues can modify the activation in early visual regions. Another possibility is that the behavioral facilitatory effects represent higher level cognitive and decision processes. Behavioral cueing studies have employed signal detection theory approaches to make claims about the locus of the lexical cueing effect, associating the sensitivity measure with lower level perceptual modifications and the bias measure with higher level, decision processes. We argue that this approach is misleading and has not been supported by studies directly looking at the neural correlates of biases, as defined by signal detection theory (Linares et al., 2019; Morgan et al., 1990; Ratcliff et al., 1989). Namely, while we recognize the value in distinguishing shifts in bias from shifts in sensitivity in terms of behavioral consequences, we argue that both can be accompanied by shifts in neural activation in perceptual regions, and are therefore not well suited for making claims about the locus of the effects.

Adopting a more direct approach to examining the neural loci of activation, studies investigating the scope of semantic representations in the brain are fairly consistent in their findings, whereas cueing studies with visual targets are divided on the topic. M/EEG cueing studies have reported the modulation of processes in the early visual cortex as soon as 100 ms post target presentation (Boutonnet & Lupyan, 2015; Hirschfeld et al., 2011; Landau et al., 2010; Noorman et al., 2018). These results follow the reports from the fMRI literature examining semantic processing in the absence of the visual target, showing that linguistic effect on perception can be observed at the early stages of visual processing. They are also in agreement with studies investigating more general, nonlinguistic feature-based top-down influences on perception (Dunovan et al., 2014; Gong & Liu, 2019; Liu et al., 2007; Maunsell & Treue, 2006; Saenz et al., 2002; Summerfield & Egner, 2016).

However, evidence coming from fMRI cueing studies stands in contrast to both fMRI studies on semantic representations in the brain and M/EEG cueing studies with a visual target. One reason for that could be that these studies have been looking for the wrong or incomplete signature of the effect. The few studies done on the topic so far have reported only univariate fMRI analyses, whereas we propose that multivariate approaches might be more appropriate to shed light on the more subtle neural signatures of language–perception interaction. Specifically, we propose that linguistic cueing studies need to directly test the existence of the sharpening account of neural activation, proposed by nonlinguistic cueing studies (Kok & de Lange, 2014; Kok et al., 2012).

Furthermore, a clearer picture is needed on the condition that would lead to inhibitory cueing effects. Current accounts explain the inhibition by claiming a lack of feature-based overlap between cue and target; however, some studies also bring up the issue of task difficulty. The relationship between the (lack of) feature overlap and inhibition or enhancement effects would therefore need to be directly tested in an experimental setting while keeping the task constant.

Additionally, the neural correlates and experimental settings leading toward bias or sensitivity shifts should be directly measured. This would help us understand under which circumstances top-down lexical influence biases our perception toward what is labeled, and under which circumstances it aids in our ability to discriminate between labeled and nonlabeled percepts.

Our current understanding is that the effects of language on perception rely on in-the-moment imagery evoked by words, which then neurally coactivates the perceptual network, and subsequently can affect how the following sensory stimuli are perceived. It would therefore be informative to see whether people with aphantasia (i.e., inability to create mental visual imagery) are resistant to linguistic top-down influences on perception.

Furthermore, any discussion about the interaction between language and perception would be incomplete without also considering the reverse process—how visual perception can influence our language comprehension and production. To our knowledge, this topic is currently underrepresented, yet it has the potential to provide deeper insights into the bidirectional flow of information between top-down and bottom-up processes between these two domains. And, as discussed in the previous paragraph, the process of perception influencing language encoding might depend on the immediate, silent naming of the perceived stimuli. This silent verbalization (i.e., verbal imagery) that might drive the conceptualization of the percept, could then neurally coactivate the language network, affecting how subsequent verbal information is received and interpreted. While several studies examined how simultaneously presented spatial and verbal information affects semantic processing (Berndt et al., 2019; Kaschak et al., 2005; Zwaan & Yaxley, 2003a, 2003b), future studies should look at the sequential effects of perception on language by using perceptual cues followed by verbal targets, in order to get a better picture of the perception–language influence.

Finally, while we have briefly mentioned several studies looking into linguistic effects on other perceptual modalities such as audition and olfaction in this review, more research needs to be done on perceptual modalities beyond vision, and their susceptibility to linguistic top-down effects.

Further research addressing these remaining issues can delineate which cognitive and neural mechanisms support linguistic influences on perception (and vice versa) and help connect those mechanisms to more general processes underlying top-down expectation and attention effects on perception.

References

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2017). What does semantic tiling of the cortex tell us about semantics? *Neuropsychologia*, *105*, 18–38. https://doi.org/10.1016/j.neuropsychologia.2017.04.011

Berndt, E., Dudschig, C., Miller, J., & Kaup, B. (2019). A replication attempt of hemispheric differences in semantic-relatedness judgments (Zwaan & Yaxley, 2003). *Acta Psychologica*, *198*, Article 102871. https://doi.org/10.1016/j.actpsy.2019.102871

Bidet-Ildei, C., Sparrow, L., & Coello, Y. (2011). Reading action word affects the visual perception of biological motion. *Acta Psychologica*, *137*(3), 330–334. https://doi.org/10.1016/j.actpsy.2011.04.001

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

Borghi, A. M., & Cimatti, F. (2009). Words as tools and the problem of abstract word meanings. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *31*, 2304–2309. https://escholarship.org/uc/item/58m9n8rp

Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object Recognition. *The Journal of Neuroscience*, *35*(25), 9329–9335. https://doi.org/10.1523/JNEUROSCI.5111-14.2015

Chasteen, A. L., Burdzy, D. C., & Pratt, J. (2010). Thinking of God moves attention. *Neuropsychologia*, *48*(2), 627–630. https://doi.org/10.1016/j.neuropsychologia.2009.09.029

de Araujo, I. E., Rolls, E. T., Velazco, M. I., Margot, C., & Cayeux, I. (2005). Cognitive modulation of olfactory processing. *Neuron*, *46*(4), 671–679. https://doi.org/10.1016/j.neuron.2005.04.021

de Gardelle, V., Waszczuk, M., Egner, T., & Summerfield, C. (2013). Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cerebral Cortex*, *23*(9), 2235–2244. https://doi.org/10.1093/cercor/bhs211

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(24), 13494–13499. https://doi.org/10.1073/pnas.93.24.13494

Dils, A. T., & Boroditsky, L. (2010). Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(37), 16396–16400. https://doi.org/10.1073/pnas.1009438107

Dove, G. (2020). More than a scaffold: Language is a neuroenhancement. *Cognitive Neuropsychology*, *37*(5–6), 288–311. https://doi.org/10.1080/02643294.2019.1637338

Drivonikou, G. V., Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., & Davies, I. R. L. (2007). Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(3), 1097–1102. https://doi.org/10.1073/pnas.0610132104

Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., & Kaup, B. (2012). From top to bottom: Spatial shifts of attention caused by linguistic stimuli. *Cognitive Processing*, *13*(Suppl. 1), 151–154. https://doi.org/10.1007/s10339-012-0480-x

Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., & Kaup, B. (2013). Reading "sun" and looking up: The influence of language on saccadic eye movements in the vertical dimension. *PLOS ONE*, *8*(2), Article e56872. https://doi.org/10.1371/journal.pone.0056872

Dunn, B., Kamide, Y., & Scheepers, C. (2014). Hearing 'moon' and looking up: Word-related spatial associations facilitate saccades to congruent locations. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*, 433–438. https://escholarship.org/content/qt78z895zd/qt78z895zd.pdf

Dunovan, K. E., Tremel, J. J., & Wheeler, M. E. (2014). Prior probability and feature predictability interactively bias perceptual decisions. *Neuropsychologia*, *61*, 210–221. https://doi.org/10.1016/j.neuropsychologia.2014.06.024

Edmiston, P., & Lupyan, G. (2015). What makes words special? Words as unmotivated cues. *Cognition*, *143*, 93–100. https://doi.org/10.1016/j.cognition.2015.06.008

Eger, E., Henson, R. N., Driver, J., & Dolan, R. J. (2007). Mechanisms of top-down facilitation in perception of visual objects studied by FMRI. *Cerebral Cortex*, *17*(9), 2123–2133. https://doi.org/10.1093/cercor/bhl119

Estes, Z., Verges, M., & Adelman, J. S. (2015). Words, objects, and locations: Perceptual matching explains spatial interference and facilitation. *Journal of Memory and Language*, *84*, 167–189. https://doi.org/10.1016/j.jml.2015.06.002

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science*, *19*(2), 93–97. https://doi.org/10.1111/j.1467-9280.2008.02051.x

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Behavioral and Brain Sciences*, *39*, Article e229. https://doi.org/10.1017/S0140525X15000965

Forder, L., & Lupyan, G. (2019). Hearing words changes color perception: Facilitation of color discrimination by verbal and visual cues. *Journal of Experimental Psychology: General*, *148*(7), 1105–1123. https://doi.org/10.1037/xge0000560

Forder, L., Taylor, O., Mankin, H., Scott, R. B., & Franklin, A. (2016). Colour terms affect detection of colour and colour-associated objects suppressed from visual awareness. *PLOS ONE*, *11*(3), Article e0152212. https://doi.org/10.1371/journal.pone.0152212

Francken, J. C., Kok, P., Hagoort, P., & de Lange, F. P. (2015). The behavioral and neural effects of language on motion perception. *Journal of Cognitive Neuroscience*, *27*(1), 175–184. https://doi.org/10.1162/jocn_a_00682

Francken, J. C., Meijs, E. L., Hagoort, P., van Gaal, S., & de Lange, F. P. (2015). Exploring the automaticity of language-perception interactions: Effects of attention and awareness. *Scientific Reports*, *5*(1), Article 17725. https://doi.org/10.1038/srep17725

Georgeson, M. (2012). Sensory, perceptual and response biases: The criterion concept in perception. *Journal of Vision*, *12*(9), Article 1392. https://doi.org/10.1167/12.9.1392

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 489–494. https://doi.org/10.1073/pnas.0509868103

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350–363. https://doi.org/10.1038/nrn3476

Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron*, *54*(5), 677–696. https://doi.org/10.1016/j.neuron.2007.05.019

Gong, M., & Liu, T. (2019). *Biased neural coding of feature-based attention in human brain*. bioRxiv. https://doi.org/10.1101/688226

Gotts, S. J., Milleville, S. C., & Martin, A. (2021). Enhanced inter-regional coupling of neural responses and repetition suppression provide separate contributions to long-term behavioral priming. *Communications Biology*, *4*(1), Article 487. https://doi.org/10.1038/s42003-021-02002-7

Gozli, D. G., Chasteen, A. L., & Pratt, J. (2013). The cost and benefit of implicit spatial cues for visual attention. *Journal of Experimental Psychology: General*, *142*(4), 1028–1046. https://doi.org/10.1037/a0030362

Gozli, D. G., Pratt, J., Martin, K. Z., & Chasteen, A. L. (2016). Implied spatial meaning and visuospatial bias: Conceptual processing influences processing of visual targets and distractors. *PLOS ONE*, *11*(3), Article e0150928. https://doi.org/10.1371/journal.pone.0150928

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Hauk, O., & Tschentscher, N. (2013). The body of evidence: What can neuroscience tell us about embodied semantics? *Frontiers in Psychology*, *4*, Article 50. https://doi.org/10.3389/fpsyg.2013.00050

Henson, R. N. (2003). Neuroimaging studies of priming. *Progress in Neurobiology*, *70*(1), 53–81. https://doi.org/10.1016/S0301-0082(03)00086-8

Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*(5456), 1269–1272. https://doi.org/10.1126/science.287.5456.1269

Herz, R. S., & von Clef, J. (2001). The influence of verbal labeling on the perception of odors: Evidence for olfactory illusions? *Perception*, *30*(3), 381–391. https://doi.org/10.1068/p3179

Hirschfeld, G., Zwitserlood, P., & Dobel, C. (2011). Effects of language comprehension on visual processing—MEG dissociates early perceptual and late N400 effects. *Brain and Language*, *116*(2), 91–96. https://doi.org/10.1016/j.bandl.2010.07.002

Hoenig, K., Sim, E.-J., Bochev, V., Herrnberger, B., & Kiefer, M. (2008). Conceptual flexibility in the human brain: Dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *Journal of Cognitive Neuroscience*, *20*(10), 1799–1814. https://doi.org/10.1162/jocn.2008.20123

Holmes, K. J., & Wolff, P. (2012). Does categorical perception in the left hemisphere depend on language? *Journal of Experimental Psychology: General*, *141*(3), 439–443. https://doi.org/10.1037/a0027289

Hotaling, J. M., Cohen, A. L., Shiffrin, R. M., & Busemeyer, J. R. (2015). The dilution effect and information integration in perceptual decision making. *PLOS ONE*, *10*(9), Article e0138481. https://doi.org/10.1371/journal.pone.0138481

Huettig, F., Olivers, C. N., & Hartsuiker, R. J. (2011). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica*, *137*(2), 138–150. https://doi.org/10.1016/j.actpsy.2010.07.013

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458. https://doi.org/10.1038/nature17637

James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). The effects of visual object priming on brain activation before and after recognition. *Current Biology*, *10*(17), 1017–1024. https://doi.org/10.1016/S0960-9822(00)00655-2

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*(3), B79–B89. https://doi.org/10.1016/j.cognition.2004.06.005

Kemmerer, D. (2023). Grounded cognition entails linguistic relativity: A neglected implication of a major semantic theory. *Topics in Cognitive Science*, *15*(4), 615–647. https://doi.org/10.1111/tops.12628

Kerzel, D., Zarian, L., & Souto, D. (2009). Involuntary cueing effects on accuracy measures: Stimulus and task dependence. *Journal of Vision*, *9*(11), Article 16. https://doi.org/10.1167/9.11.16

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *48*(7), 805–825. https://doi.org/10.1016/j.cortex.2011.04.006

Kok, P., & de Lange, F. P. (2014). Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology*, *24*(13), 1531–1535. https://doi.org/10.1016/j.cub.2014.05.042

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Landau, A. N., Aziz-Zadeh, L., & Ivry, R. B. (2010). The influence of language on perception: Listening to sentences about faces affects the perception of faces. *The Journal of Neuroscience*, *30*(45), 15254–15261. https://doi.org/10.1523/JNEUROSCI.2046-10.2010

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America, A, Optics, Image Science & Vision*, *20*(7), 1434–1448. https://doi.org/10.1364/JOSAA.20.001434

Linares, D., Aguilar-Lleyda, D., & López-Moliner, J. (2019). Decoupling sensory from decisional choice biases in perceptual decision making. *eLife*, *8*, Article e43994. https://doi.org/10.7554/eLife.43994

Liu, T., Stevens, S. T., & Carrasco, M. (2007). Comparing the time course and efficacy of spatial and feature-based attention. *Vision Research*, *47*(1), 108–113. https://doi.org/10.1016/j.visres.2006.09.017

Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, *3*(2), 273–302. https://doi.org/10.1111/j.1756-8765.2010.01106.x

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, *3*, Article 54. https://doi.org/10.3389/fpsyg.2012.00054

Lupyan, G., Abdel Rahman, R., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in Cognitive Sciences*, *24*(11), 930–944. https://doi.org/10.1016/j.tics.2020.08.005

Lupyan, G., & Spivey, M. J. (2010a). Making the invisible visible: Verbal but not visual cues enhance visual detection. *PLOS ONE*, *5*(7), Article e11452. https://doi.org/10.1371/journal.pone.0011452

Lupyan, G., & Spivey, M. J. (2010b). Redundant spoken labels facilitate perception of multiple items. *Attention, Perception, & Psychophysics*, *72*(8), 2236–2253. https://doi.org/10.3758/BF03196698

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*, *141*(1), 170–186. https://doi.org/10.1037/a0024904

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(35), 14196–14201. https://doi.org/10.1073/pnas.1303312110

Manescu, S., Frasnelli, J., Lepore, F., & Djordjevic, J. (2014). Now you like me, now you don't: Impact of labels on odor perception. *Chemical Senses*, *39*(2), 167–175. https://doi.org/10.1093/chemse/bjt066

Martens, U., & Gruber, T. (2012). Sharpening and formation: Two distinct neuronal mechanisms of repetition priming. *European Journal of Neuroscience*, *36*(7), 2989–2995. https://doi.org/10.1111/j.1460-9568.2012.08222.x

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, *270*(5233), 102–105. https://doi.org/10.1126/science.270.5233.102

Mathôt, S., Grainger, J., & Strijkers, K. (2017). Pupillary responses to words that convey a sense of brightness or darkness. *Psychological Science*, *28*(8), 1116–1124. https://doi.org/10.1177/0956797617702699

Mathôt, S., Sundermann, L., & van Rijn, H. (2019). *The effect of semantic brightness on pupil size: A replication with Dutch words*. bioRxiv. https://doi.org/10.1101/689265

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. https://doi.org/10.1016/j.tins.2006.04.001

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science*, *18*(11), 1007–1013. https://doi.org/10.1111/j.1467-9280.2007.02016.x

Mo, L., Xu, G., Kay, P., & Tan, L.-H. (2011). Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(34), 14026–14030. https://doi.org/10.1073/pnas.1111860108

Morgan, M. J., Hole, G. J., & Glennerster, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research*, *30*(11), 1793–1810. https://doi.org/10.1016/0042-6989(90)90160-M

Müller, N. G., Strumpf, H., Scholz, M., Baier, B., & Melloni, L. (2013). Repetition suppression versus enhancement—It's quantity that matters. *Cerebral Cortex*, *23*(2), 315–322. https://doi.org/10.1093/cercor/bhs009

Nijhof, A. D., & Willems, R. M. (2015). Simulating fiction: Individual differences in literature comprehension revealed with FMRI. *PLOS ONE*, *10*(2), Article e0116492. https://doi.org/10.1371/journal.pone.0116492

Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, *13*(2), 248–277. https://doi.org/10.1016/0010-0285(81)90010-4

Noorman, S., Neville, D. A., & Simanova, I. (2018). Words affect visual perception by activating object shape representations. *Scientific Reports*, *8*(1), Article 14156. https://doi.org/10.1038/s41598-018-32483-2

Okamoto, M., Wada, Y., Yamaguchi, Y., Kimura, A., Dan, H., Masuda, T., Singh, A. K., Clowney, L., & Dan, I. (2009). Influences of food-name labels on perceived tastes. *Chemical Senses*, *34*(3), 187–194. https://doi.org/10.1093/chemse/bjn075

Orgs, G., Lange, K., Dombrowski, J.-H., & Heil, M. (2006). Conceptual priming for environmental sounds and words: An ERP study. *Brain and Cognition*, *62*(3), 267–272. https://doi.org/10.1016/j.bandc.2006.05.003

Ostarek, M., & Huettig, F. (2017). Spoken words can make the invisible visible—Testing the involvement of low-level visual representations in spoken word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(3), 499–508. https://doi.org/10.1037/xhp0000313

Ostarek, M., & Vigliocco, G. (2017). Reading sky and seeing a cloud: On the relevance of events for perceptual simulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(4), 579–590. https://doi.org/10.1037/xlm0000318

Pavan, A., Skujevskis, M., & Baggio, G. (2013). Motion words selectively modulate direction discrimination sensitivity for threshold motion. *Frontiers in Human Neuroscience*, *7*, Article 134. https://doi.org/10.3389/fnhum.2013.00134

Pelekanos, V., & Moutoussis, K. (2011). The effect of language on visual contrast sensitivity. *Perception*, *40*(12), 1402–1412. https://doi.org/10.1068/p7010

Pinto, Y., van Gaal, S., de Lange, F. P., Lamme, V. A. F., & Seth, A. K. (2015). Expectations accelerate entry of visual stimuli into awareness. *Journal of Vision*, *15*(8), Article 13. https://doi.org/10.1167/15.8.13

Pirog Revill, K., Aslin, R. N., Tanenhaus, M. K., & Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(35), 13111–13115. https://doi.org/10.1073/pnas.0807054105

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576–582. https://doi.org/10.1038/nrn1706

Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, *17*(9), 458–470. https://doi.org/10.1016/j.tics.2013.06.004

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, *11*(5), 351–360. https://doi.org/10.1038/nrn2811

Puri, A. M., Wojciulik, E., & Ranganath, C. (2009). Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Research*, *1301*, 89–99. https://doi.org/10.1016/j.brainres.2009.08.085

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341–365. https://doi.org/10.1017/S0140525X99002022

Ratcliff, R., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 378–387. https://doi.org/10.1037/0278-7393.15.3.378

Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences*, *13*(10), 439–446. https://doi.org/10.1016/j.tics.2009.07.001

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, *27*(5), 767–780. https://doi.org/10.1207/s15516709cog2705_4

Rueschemeyer, S.-A., Glenberg, A. M., Kaschak, M. P., Mueller, K., & Friederici, A. D. (2010). Top-down and bottom-up contributions to understanding sentences describing objects in motion. *Frontiers in Psychology*, *1*, Article 183. https://doi.org/10.3389/fpsyg.2010.00183

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience*, *5*(7), 631–632. https://doi.org/10.1038/nn876

Saygin, A. P., McCullough, S., Alac, M., & Emmorey, K. (2010). Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *Journal of Cognitive Neuroscience*, *22*(11), 2480–2490. https://doi.org/10.1162/jocn.2009.21388

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia*, *45*(12), 2802–2810. https://doi.org/10.1016/j.neuropsychologia.2007.05.002

Slivac, K., & Flecken, M. (2023). Linguistic priors for perception. *Topics in Cognitive Science*, *15*(4), 657–661. https://doi.org/10.1111/tops.12672

Slivac, K., Hervais-Adelman, A., Hagoort, P., & Flecken, M. (2021). Linguistic labels cue biological motion perception and misperception. *Scientific Reports*, *11*(1), Article 17239. https://doi.org/10.1038/s41598-021-96649-1

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *The Journal of Neuroscience*, *32*(25), 8443–8453. https://doi.org/10.1523/JNEUROSCI.5069-11.2012

Sorokowska, A., Albrecht, E., & Hummel, T. (2015). Reading first or smelling first? Effects of presentation order on odor identification. *Attention, Perception, & Psychophysics*, *77*, 731–736. https://doi.org/10.3758/s13414-014-0811-3

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326

Stein, T., & Peelen, M. V. (2015). Content-specific expectations enhance stimulus detectability by increasing perceptual sensitivity. *Journal of Experimental Psychology: General*, *144*(6), 1089–1104. https://doi.org/10.1037/xge0000109

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745–756. https://doi.org/10.1038/nrn3838

Summerfield, C., & Egner, T. (2016). Feature-based attention and feature-based expectation. *Trends in Cognitive Sciences*, *20*(6), 401–404. https://doi.org/10.1016/j.tics.2016.03.008

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004–1006. https://doi.org/10.1038/nn.2163

Taylor, J. E. T., Lam, T. K., Chasteen, A. L., & Pratt, J. (2015). Bow your head in shame, or, hold your head up with pride: Semantic processing of self-esteem concepts orients attention vertically. *PLOS ONE*, *10*(9), Article e0137704. https://doi.org/10.1371/journal.pone.0137704

Tillas, A. (2015). Language as grist to the mill of cognition. *Cognitive Processing*, *16*(3), 219–243. https://doi.org/10.1007/s10339-015-0656-2

Turk-Browne, N. B., Yi, D.-J., Leber, A. B., & Chun, M. M. (2007). Visual quality determines the direction of neural repetition effects. *Cerebral Cortex*, *17*(2), 425–433. https://doi.org/10.1093/cercor/bhj159

Verges, M., & Duffy, S. (2009). Spatial representations elicit dual-coding effects in mental imagery. *Cognitive Science*, *33*(6), 1157–1172. https://doi.org/10.1111/j.1551-6709.2009.01038.x

Wallentin, M., Nielsen, A. H., Vuust, P., Dohn, A., Roepstorff, A., & Lund, T. E. (2011). BOLD response to motion verbs in left posterior middle temporal gyrus during story comprehension. *Brain and Language*, *119*(3), 221–225. https://doi.org/10.1016/j.bandl.2011.04.006

Wallentin, M., Weed, E., Østergaard, L., Mouridsen, K., & Roepstorff, A. (2008). Accessing the mental space-Spatial working memory processes for language and vision overlap in precuneus. *Human Brain Mapping*, *29*(5), 524–532. https://doi.org/10.1002/hbm.20413

Weger, U. W., & Pratt, J. (2008). Time flies like an arrow: Space–time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin & Review*, *15*(2), 426–430. https://doi.org/10.3758/PBR.15.2.426

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*(2), 227–233. https://doi.org/10.1016/S0959-4388(98)80144-X

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal detection measures cannot distinguish perceptual biases from response biases. *Perception*, *44*(3), 289–300. https://doi.org/10.1068/p7908

Witzel, C., & Gegenfurtner, K. R. (2011). Is there a lateralized category effect for color? *Journal of Vision*, *11*(12), Article 16. https://doi.org/10.1167/11.12.16

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 253–265. https://doi.org/10.1002/wcs.104

Yee, E., Ahmed, S. Z., & Thompson-Schill, S. L. (2012). Colorless green ideas (can) prime furiously. *Psychological Science*, *23*(4), 364–369. https://doi.org/10.1177/0956797611430691

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review*, *23*(4), 1015–1027. https://doi.org/10.3758/s13423-015-0948-7

Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, *9*(1), Article 4288. https://doi.org/10.1038/s41467-018-06752-7

Zhang, K., & Sejnowski, T. J. (1999). Neuronal tuning: To sharpen or broaden? *Neural Computation*, *11*(1), 75–84. https://doi.org/10.1162/089976699300016809

Zhou, K., Mo, L., Kay, P., Kwok, V. P. Y., Ip, T. N. M., & Tan, L. H. (2010). Newly trained lexical categories produce lateralized categorical perception of color. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(22), 9974–9978. https://doi.org/10.1073/pnas.1005669107

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171. https://doi.org/10.1111/1467-9280.00430

Zwaan, R. A., & Yaxley, R. H. (2003a). Hemispheric differences in semantic-relatedness judgments. *Cognition*, *87*(3), B79–B86. https://doi.org/10.1016/S0010-0277(02)00235-4

Zwaan, R. A., & Yaxley, R. H. (2003b). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, *10*(4), 954–958. https://doi.org/10.3758/BF03196557

Received February 17, 2023
Revision received July 15, 2024
Accepted December 31, 2024 ■