

## You heard it here first: Readers mentally simulate described sounds

Tad T. Brunyé<sup>a,c,\*</sup>, Tali Ditman<sup>b,c</sup>, Caroline R. Mahoney<sup>a,c</sup>, Eliza K. Walters<sup>c</sup>, Holly A. Taylor<sup>c</sup>

<sup>a</sup> U.S. Army NSRDEC, Cognitive Science, Natick, MA, USA

<sup>b</sup> Harvard Medical School and Massachusetts General Hospital, Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA

<sup>c</sup> Tufts University, Department of Psychology, Medford, MA, USA

### ARTICLE INFO

#### Article history:

Received 26 April 2010

Received in revised form 18 June 2010

Accepted 21 June 2010

Available online 10 July 2010

#### PsychInfo Classification:

2343

#### Keywords:

Embodied cognition

Auditory imagery

Language

### ABSTRACT

The present experiments examined whether readers spontaneously simulate implied auditory elements of sentences. Participants read sentences that implicitly conveyed details that could provoke auditory imagery (e.g., The engine clattered as the truck driver warmed up his rig.), and then performed an unrelated sound categorization task during which they classified sounds as real (occurring in the world) or fake (computer generated). In Experiment 1 these two tasks were performed in sequence; in Experiment 2 they were separated into three experimental blocks to rule out the possibility that readers strategically formed auditory imagery as a result of task demands. In both studies, readers were faster to correctly categorize sounds as 'real' when the sounds had been implied by a preceding sentence. These results suggest that readers mentally simulate the implied auditory characteristics of sentences, even in the absence of tasks that promote mental simulation. Mentally simulating described events is not limited to visual and action-based modalities, further demonstrating the multimodal nature of the perceptual symbols spontaneously activated during reading.

Published by Elsevier B.V.

Understanding language involves both the comprehension of individual words and propositions and the development of rich multidimensional situation models that code for information regarding space, time, characters, actions, objects, emotions, causality and expectations (Ditman, Holcomb, & Kuperberg, 2008; Gernsbacher, 1990; Gernsbacher & Foertsch, 1999; Graesser, Gernsbacher, & Goldman, 2003; Zwaan, 1999; Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). A growing body of work has also suggested that readers routinely activate sensorimotor information during language processing, and that such activations may play an important role in successful understanding (Bergen, Lindsay, Matlock, & Narayanan, 2007; Kaschak & Glenberg, 2000; Zwaan, 2004). Indeed readers activate visual and motor representations during language comprehension, and brain regions responsible for these representations tend to overlap with those activated during real-world perception and action (e.g., Brunyé, Ditman, Mahoney, Augustyn, & Taylor, 2009; Ditman, Brunyé, Mahoney, & Taylor, 2010; Glenberg & Kaschak, 2002; Pulvermüller, 2005; Richardson, Spivey, Barsalou, & McRae, 2003; Ruby & Decety, 2001). In general, this growing body of evidence has been taken as strong support for theories positing the importance of embodied mental simulations in comprehending language (Barsalou, 2005, 2008; Fischer & Zwaan, 2008; Glenberg, 1997). The goal of the present studies is to extend previous work on embodied

mental simulations to the domain of auditory processing. Specifically, we are interested in whether evidence for mental simulations can be found in a non-dominant sensory modality such as audition, or if the mental representation of sound is better accounted for by theories positing amodal propositions as the primary mechanism underlying language comprehension and memory (i.e., Anderson, 1990; Fodor, 1975, 2000; Kintsch, 1988; Pylyshyn, 2002). To this end, Experiment 1 examines whether readers perform auditory mental simulations during language comprehension; if this is the case, it provides strong evidence that theories of embodied language comprehension are not limited to motor or visual representations (i.e., Posner, Nissen, & Klein, 1976) and can be extended to a larger range of sensorimotor representation. Experiment 2 further examines whether these mental simulations occur spontaneously and in the absence of experimental tasks that may promote the strategic use of auditory imagery.

### 1. Sensorimotor simulation during reading

Recent theories of language comprehension posit that readers construct experiential simulations of described visual and motor information, grounding linguistic meaning in bodily senses and activities (i.e., Barsalou, 1999; Fincher-Kiefer, 2001; Glenberg, 1997; Lakoff, 1987; Zwaan, 2004). This theoretical stance is in direct contrast to more traditional amodal theories of cognition in general, and language comprehension in particular (i.e., Collins & Quillian, 1969; Fodor, 1975, 2000; Kintsch, 1988; Pylyshyn, 1984). There are four primary research categories lending support for mental simulations during reading (cf., Zwaan, 2004). First, when

\* Corresponding author. Tufts University, Department of Psychology, 490 Boston Ave, Medford, MA 02155, USA. Tel.: +1 617 306 6262.

E-mail address: [tbrunye@alumni.tufts.edu](mailto:tbrunye@alumni.tufts.edu) (T.T. Brunyé).

participants read sentences they quickly activate visual representations corresponding to an unfolding narrative; for instance, readers mentally simulate the shape, visibility and orientation of described objects (Stanfield & Zwaan, 2001; Zwaan et al., 2002; Zwaan & Yaxley, 2003; Yaxley & Zwaan, 2007). Second, words activate motor representations that can affect subsequent motor activity. For instance, readers perform manual responses faster when they are congruent rather than incongruent with an action described in a sentence (i.e., the Action Compatibility Effect; Glenberg & Kaschak, 2002; Tucker & Ellis, 2004; Zwaan & Taylor, 2006). Readers also make eye movements that correspond with those that might take place in the situation being described (Spivey, Richardson, Tyler, & Young, 2000), and alter hand aperture in grasping tasks as a function of inferred size gathered from linguistic stimuli (i.e., grape versus apple; Glover, Rosenbaum, Graham, & Dixon, 2004). Third, there is evidence that perceptual information associated with currently unfolding events is more accessible than information outside of those events; for instance, participants verify the presence or object properties faster when they are spatially closer to a protagonist's present location (e.g., Borghi, Glenberg, & Kaschak, 2004; Horton & Rapp, 2003; Morrow et al., 1987). Finally, a growing body of neuroimaging data demonstrates that readers activate brain regions that closely align with those activated during real-world perception and action (Martin & Chao, 2001; Pulvermüller, 2005, 2008; Tettamanti et al., 2005); for instance, listening to action related sentences activates the areas of the premotor cortex where the specific actions are coded (Pulvermüller, 2005; Tettamanti et al., 2005).

Although these findings provide convincing evidence that readers routinely and immediately form mental simulations of described events, some contemporary language theories express the importance of amodal symbols in representing language (e.g., Landauer, 2002; Landauer, McNamara, Dennis, & Kintsch, 2007). Indeed there is compelling evidence that symbolic non-perceptual frameworks, such as those underlying some computational models (i.e., HAL, LSA), can account for a range of human performance (Burgess & Lund, 1997; Landauer, Laham, & Foltz, 1998). To further parse the relative contributions of amodal and modal symbols during language comprehension, research must consider the full range of human perceptual experience (i.e., Zwaan, 2004). However, current research examining the role of perceptual symbols in language comprehension is largely limited to visual imagery and motor movement. Most studies examining the interaction between language comprehension and mental simulation have used visual stimuli and motor responses to demonstrate how words can guide and constrain perception and action (e.g., Brunyé et al., 2009; Chambers, Tanenhaus, & Magnuson, 2004; Kaschak et al., 2005; Zwaan et al., 2002; Stanfield & Zwaan, 2001). To the extent that readers develop comprehensive multimodal mental representations of described events, and mental simulation is an important component of comprehension, then such simulations should also be expected to code for the broad range of human sensory activations potentially imparted by text, even in non-dominant sensory modalities. The present research aims to partially fill this research gap by examining whether readers mentally simulate described sounds.

## 2. Auditory mental simulations

To our knowledge, only one study has specifically examined whether readers mentally simulate auditory properties described in a text (Kaschak, Zwaan, Aveyard, & Yaxley, 2006). Kaschak and colleagues had participants read sentences that emphasized auditory aspects of described events while having them listen to white noise that either matched (e.g., white noise getting louder) or mismatched (e.g., white noise getting quieter) a described motion (e.g., Waving frantically, the lost hiker heard the chopper head toward him.). The authors found that participants made sensibility judgments (i.e., yes/

no judgments of whether a sentence 'makes sense') more slowly when the auditory and linguistic cues matched; that is, conflict occurred when auditory mechanisms had to process both the presented sound and the implied sound in the sentence.

These results provide seminal support for the notion that readers mentally simulate sounds during reading. However, they cannot rule out the possibility that readers only simulate the implied motion of described sounds (i.e., white noise moving up or down, or moving towards or away from a reader), rather than simulating the specific sound being described (i.e., a chainsaw moving up or down). Further, they cannot rule out the possibility that readers only mentally simulate sounds under experimental task demands. Because the auditory and linguistic stimuli were inherently related and simultaneously presented, conflict between these representations may only arise when participants deem both stimuli as important for task performance. Indeed some have argued that experimental designs promoting a meaningful overlap between language and perception may promote the strategic use of imagery that does not occur spontaneously during more natural language comprehension (cf. Fiedler, 2009; Machery, 2007; van Dantzig, Pecher, Zeelenberg, & Barsalou, 2008). It is therefore critical to demonstrate that readers mentally simulate specific sounds, and do so without the explicit introduction of auditory information during reading. Our first experiment explores these issues by testing whether reading activates specific sounds during a task that does not unnecessarily draw attention to auditory simulation.

Though very limited work has examined the simulation of sounds during reading, some evidence may be derived from studies examining Auditory Imagery Experiences (AIEs). AIEs occur when readers reactivate memories of a character's voice and its unique rate, gender, prosody, timbre and pitch. Recent research has demonstrated that such re-experiencing of voices during reading appears to be quite common and contingent upon sufficient prior exposure to the voice (Alexander & Nygaard, 2008; Kurby, Magliano, & Rapp, 2009), suggesting that readers may indeed perform auditory simulations that code for specific and distinctive auditory components of described events.

## 3. Experiment 1

To determine whether readers spontaneously simulate specific sounds during reading, participants read sentences that included an activity associated with a specific sound (e.g., The engine clattered as the truck driver warmed up his rig.) and then performed an ostensibly unrelated categorization task in which they decided whether a probe sound was real or computer-generated; on one quarter of the trials the sound was real and matched the sound implied by the preceding sentence. If readers spontaneously form auditory imagery during sentence comprehension, then we predict that probe categorization should be faster when a real sound specifically matches rather than mismatches the sound implied by a preceding sentence. This hypothesis is based on research demonstrating that sensorimotor activation during language comprehension can facilitate performance on subsequent visual- or action-based verification tasks using congruent rather than incongruent stimulus properties (e.g., Glenberg & Kaschak, 2002; Glenberg et al., 2008; Stanfield & Zwaan, 2001; Yaxley & Zwaan, 2007). Such a result would provide strong evidence that readers form auditory simulations of implied sounds, and also that these can occur even with seemingly unrelated linguistic and auditory stimuli; it would also support the notion that conceptual representations are built from the activation of perceptual symbols that involve the full range of human sensory capabilities (i.e., Perceptual Symbols Theory; Barsalou, 1999). The amodal propositional framework provides a contrasting hypothesis; in this framework, readers represent the relationships between units of language in a symbolic nature without necessarily grounding them in traces of perception or action (i.e., Collins & Quillian, 1969; Fodor, 1975, 2000;

Kintsch, 1988; Landauer, 2002; Landauer, McNamara, Dennis, & Kintsch, 2007; Pylyshyn, 1984). Under this account, reading about events should not spontaneously reactivate experiential auditory traces, and thus should not affect performance on a sound categorization task. The present study allows for a test between these two competing hypotheses.

### 3.1. Method

#### 3.1.1. Participants and design

Eighty Tufts University undergraduates (39 males; age  $M = 20.2$ ) participated for monetary compensation. The study used a  $2(\text{Sound Type: real, fake}) \times 2(\text{Sentence Congruence: match, mismatch})$  repeated-measures design to examine time to categorize sounds (real or fake) that either matched or mismatched the sound implied in a preceding sentence. To encourage reading for comprehension and memory, we also tested memory for sentence nouns in a recognition test.

#### 3.1.2. Materials

**3.1.2.1. Sentences.** Twenty-four sentences described rich auditory experiences embedded within simple events (see Table 1). To maintain discourse focus on the implied auditory aspect of the event, all sentences began with a noun phrase preceded by a definite article in attentional focus (e.g., The campfire crackled...) and ended with a simple context (e.g., ...as the kids prepared for storytime.).

**3.1.2.2. Sounds.** Twenty-four sound files were selected from an internet sound database (<http://www.freesound.org>); each sound file was an actual recording of a continuous real-world event corresponding to the sentence stimuli (e.g., campfire crackling, horns honking, engine running, and jackhammer pounding). Sounds were normalized in amplitude to 89 dB, and standardized to 5 s in duration. To ensure identifiable sounds we made an effort to select only ordinary and high frequency sounds. Modified 'fake' sounds were developed by adding digitized sound effects to the original sound files with the Audacity audio editor (<http://audacity.sourceforge.net>).

We conducted a pilot study ( $n = 14$ ) in which participants categorized the 24 sounds as 'real' or 'fake' and then provided an open-ended written identification of the sound; results demonstrated highly accurate sound categorization (real versus fake; 98% correct categorization), and that each original 'real' sound was easily and accurately labeled (correct labeling, 94%), and modified 'fake' sounds were not (correct labeling, 15%).

**Table 1**

Example sentence and sound stimuli used in Experiments 1 and 2.

Sentence	Sound
The campfire crackled as the kids prepared for storytime. The crows sang in the trees while the couple had a picnic. The engine clattered as the truck driver warmed up his rig.	Crackling campfire Crows cawing Diesel engine starting
The jackhammer pounded the pavement where the new bank was being built. The waves crashed against the beach as the tide began to rise. The orchestra tuned their instruments as the patrons found their seats.	Jackhammer on pavement Waves on beach String orchestra
The chainsaw buzzed as it tore through the tree trunk. The crowd roared as the team scored a third goal.	Chainsaw on wood Soccer stadium cheers
The horns honked as the drivers got angry at the traffic. The siren blazed as the ambulance sped through the intersection.	Honking horns Ambulance siren
The water gushed out of the spigot into the filling bathtub. The vacuum hummed in the room where the maid was working.	Running water Vacuum on carpet

**3.1.2.3. Recognition test.** A yes/no recognition test included 48 trials, each containing either a previously read (e.g., campfire) or unread (lure) noun (e.g., bonfire). To ensure task difficulty we made an effort to select lure nouns that reasonably matched the sentence context and were at least moderately associated with learned nouns (using LSA; <http://lsa.colorado.edu>).

#### 3.1.3. Procedure

**3.1.3.1. Instructions.** Participants were told that they would be performing two tasks in sequence: reading sentences, and categorizing sounds as either real or fake. For the reading task, they were instructed to read each sentence carefully for a later memory test. For the sound categorization task, they were instructed to listen carefully to the presented sound and determine as quickly and accurately as possible whether the sound was 'real' or 'fake.' They were told that the researchers carried a microphone and recorder to dozens of locations throughout New England, and recorded things that they heard; when the researchers returned from their travels, they used audio editing software to make digitized 'fake' versions of the sounds. Participants were given four practice trials demonstrating the type of sentences they would read and the auditory distinction between real and fake sounds. Finally, for the recognition task, participants were instructed to respond to each word as 'old' or 'new' as quickly and accurately as possible.

**3.1.3.2. Reading sentences and categorizing sounds.** Each sentence was presented one at a time on the computer monitor at a rate corresponding to 250 ms per word (i.e., Rayner, 1998). After each sentence a 500 ms sequence of central fixation symbols appeared, and then a sound file was played through headphones at fixed volume (sound files were 89 dB, operating system volume set at 20% of maximum). The sound file would continue either until a participant response or 5 s had passed. The sound file was either real (original) or fake (modified), and either a match or mismatch to the preceding sentence. Participants responded to sounds by using the labeled keys F (real) and J (fake).

Four lists were created to rotate each sentence through the four sound conditions, corresponding to a  $2 \times 2$  factorial design crossing Sound Type (real, fake) with Sentence Congruence (match, mismatch). Within each list, 6 sentences were followed by a real/match sound, 6 by a real/mismatch sound, 6 by a fake/match sound, and 6 by a fake/mismatch sound. This design was intended to minimize the risk of participants realizing our intentions; indeed the majority (75%) of trials was either fake (50%; and thus had low identifiability) or real yet unrelated (25%) to the immediately preceding sentence. Within each list, sentence and sound pairs were presented in random order. After completing 24 sentences and their corresponding sound categorizations, participants began a recognition test.

**3.1.3.3. Recognition test.** The recognition test presented 48 single-word trials in random order, during each of which the participant responded whether the word was old (previously studied) or new (a novel lure) by using the C and M keys, respectively. Trials would time out and continue to the next trial after 5 s without a response.

**3.1.3.4. Questionnaire.** To assess whether participants realized the intention of the study, we had them answer three brief questions at study end: what were you asked to do in this study, what strategies did you use to help you do what you were asked, and describe what you thought was the true intention of the study. Questionnaire results indicated that all participants understood the instructions, the majority of participants reported a strategy of 'visualizing what was described' to aid sentence memory and two of the eighty participants



accurately determined our intentions and were thus removed from subsequent analysis.

### 3.2. Results

#### 3.2.1. Categorizing sounds

We examined average response times to sound categorization trials as a function of whether the sound was real or fake and whether it matched or mismatched the preceding sentence; to test our hypotheses, we were specifically interested in the difference between matching versus mismatching sounds during real sound trials. Only correct categorizations (i.e., correctly categorizing as 'real' or 'fake') were included for analysis, comprising 84% of the data set (categorization accuracy did not vary as a function of Sound Type or Sentence Congruence,  $p > .05$ ).

A 2(Sound Type: real, fake)  $\times$  2(Sentence Congruence: match, mismatch) repeated-measures analysis of variance (ANOVA) on response times revealed no effect of Sound Type,  $F(1, 77) = 1.24$ ,  $p > .05$ , but a main effect of Sentence Congruence,  $F(1, 77) = 7.07$ ,  $p < .01$ ,  $\eta^2 = .03$ .<sup>1</sup> This latter effect was qualified by an interaction between Sound Type and Sentence Congruence,  $F(1, 77) = 18.45$ ,  $p < .01$ ,  $\eta^2 = .05$ , as depicted in Fig. 1a. To directly compare the match and mismatch conditions, we conducted two paired t-tests, one in each of the two Sound Type conditions. Within the real condition, the match condition showed significantly faster response times ( $M = 2.28$  s,  $SE = .08$ ) relative to the mismatch condition ( $M = 2.69$  s,  $SE = .10$ ),  $t(77) = 4.85$ ,  $p < .01$ ,  $d = .55$ ; this effect was not replicated in the fake condition,  $t(77) = .69$ ,  $p > .05$ .

#### 3.2.2. Recognition test

Overall, recognition accuracy was moderately high ( $M = .82$ ,  $SE = .02$ ) and false alarms moderately low ( $M = .25$ ,  $SE = .01$ ), demonstrating that participants understood the task and were reading the sentences for comprehension and memory (see Table 2). Though not specifically related to our hypotheses, to test whether recognition accuracy varied as a function of our experimental conditions, we conducted a 2(Sound Type: real, fake)  $\times$  2(Sentence Congruence: match, mismatch) ANOVA, which demonstrated an interaction between Sound Type and Sentence Congruence,  $F(1, 77) = 5.52$ ,  $p < .01$ ,  $\eta^2 = .02$  (main effects non-significant,  $p$ 's  $> .05$ ). Paired t-tests demonstrated higher accuracy for matching versus mismatching sounds, but only in the real condition,  $t(77) = 3.12$ ,  $p < .01$ ,  $d = .35$  (fake condition,  $p > .05$ ). Recognition response times mirrored accuracy rates, but an ANOVA did not reveal any significant effects (all  $p$ 's  $> .05$ ).

### 3.3. Experiment 1 discussion

Experiment 1 examined whether sentences with implied auditory elements would facilitate the categorization of matching relative to mismatching sounds on an ostensibly unrelated task. The predictions borne out of the perceptual symbols framework were generally confirmed, with participants showing faster sound categorization when that sound matched versus mismatched a preceding sentence. Amodal symbols frameworks do not predict this effect. Our result provides evidence that in addition to the implied movement of sounds (Kaschak et al., 2006), readers simulate the specific sounds being described in a text, and appear to do so even in a task that was not simultaneous to, or related to, the reading task. Further, when sounds matched a preceding sentence, participants also showed higher accuracy on a recognition task, suggesting that they integrated sounds into a multi-

**Table 2**

Accuracy and response times for the recognition task in Experiments 1 and 2.

Experiment and measure	Mean	Standard error
<i>Experiment 1: accuracy (HITs) and false alarms</i>		
Real verification, matching sound	.862	.013
Real verification, mismatching sound	.809	.018
Fake verification, matching sound	.803	.019
Fake verification, mismatching sound	.822	.019
False alarm rate	.258	.011
<i>Experiment 1: response times (in seconds)</i>		
Real verification, matching sound	1.12	.032
Real verification, mismatching sound	1.19	.042
Fake verification, matching sound	1.16	.029
Fake verification, mismatching sound	1.15	.034
False alarms	1.49	.044
<i>Experiment 2: accuracy (HITs) and false alarms</i>		
Real verification, matching sound	.706	.025
Real verification, mismatching sound	.673	.025
Fake verification, matching sound	.702	.024
Fake verification, mismatching sound	.717	.024
False alarm rate	.261	.015
<i>Experiment 2: response times (in seconds)</i>		
Real verification, matching sound	1.12	.037
Real verification, mismatching sound	1.13	.041
Fake verification, matching sound	1.22	.047
Fake verification, mismatching sound	1.11	.039
False alarms	1.39	.058

modal situation model of the described events. Experiment 2 isolated the sentences from the sound categorization task to further rule out the potential influence of task demands.

## 4. Experiment 2

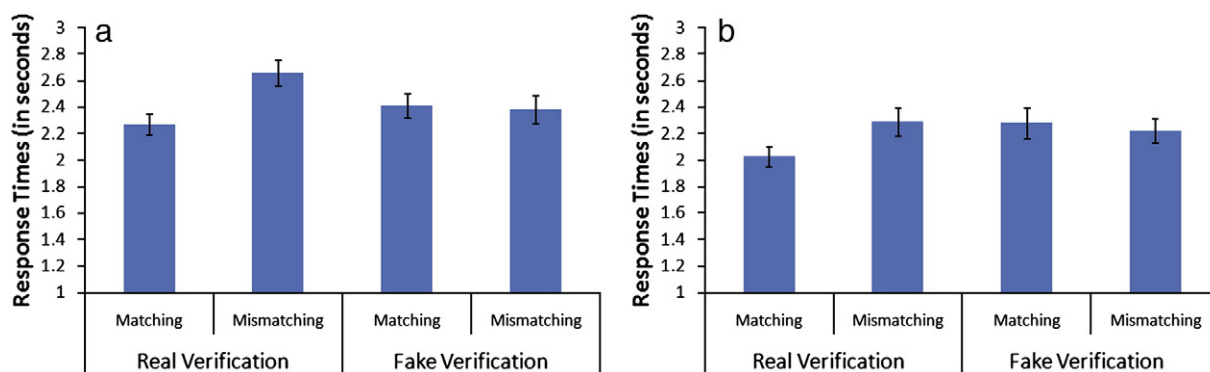
The fact that two participants in Experiment 1 accurately determined the intention of our design, points to the possibility that other participants' performance may have been affected by task demands and strategic imagery use, even without their knowledge. That is, support for perceptual symbols might only arise when tasks are designed to promote the use of mental imagery towards successful task performance (Machery, 2007). Our second experiment was designed to further rule out this potential influence. To help control for the influence of task demands, we isolated the sentence reading and sound categorization tasks by dividing the trials into three segments. During each segment, participants were presented with eight sentences in sequence (without intervening sound categorization trials) and then performed the eight corresponding sound categorization trials. In this way, the reading task was isolated from sound categorization. A replication of Experiment 1 results would increase our confidence that performance variation on the sound categorization task is a result of spontaneous auditory simulation rather than strategic imagery use (cf., Ditman et al., 2010; Machery, 2007; van Dantzig et al., 2008). Such a replication would also be in line with some recent work that has demonstrated interactions between picture viewing and language comprehension, even when the two tasks are unrelated and separated by an extended time interval (Wassenburg & Zwaan, in press).

### 4.1. Method

#### 4.1.1. Participants and design

Eighty Tufts University undergraduates (32 males; age  $M = 20$ ) participated for monetary compensation. The overall design matched that of Experiment 1.

<sup>1</sup> Given that items in our design were fully counterbalanced across all conditions, we only report subjects analyses, not items analyses (i.e., Raaijmakers, Schrijnemakers, & Gremmen, 1999).



**Fig. 1.** a. Experiment 1 mean and standard error response times during the sound categorization task, depicted as a function of whether the sound was real or fake and whether it matched or mismatched the immediately preceding sentence. b. Experiment 2 mean and standard error response times during the sound categorization task, depicted as a function of whether the sound was real or fake and whether it matched or mismatched the preceding set of sentences.

#### 4.1.2. Materials

All materials matched those used in Experiment 1.

#### 4.1.3. Procedure

**4.1.3.1. Instructions.** Instructions matched those used in Experiment 1, with the exception that participants were told that they would be performing the two tasks in three segments of eight (rather than in sequence).

**4.1.3.2. Reading sentences and categorizing sounds.** All procedures matched those of Experiment 1 with the exception of parsing the trials into three segments of sentence and sound categorization trials. As in Experiment 1, four lists were created to rotate each sentence through the four sound conditions, corresponding to a  $2 \times 2$  factorial design crossing Sound Type (real and fake) with Sentence Congruence (match and mismatch). These lists were each divided into the three segments; within each segment, 2 sentences were followed by a real/match sound, 2 by a real/mismatch sound, 2 by a fake/match sound, and 2 by a fake/mismatch sound. After all 8 sentences were presented, the eight corresponding sound categorization trials were presented. Sentences were presented in random order within each segment, as were the corresponding sound files.

**4.1.3.3. Recognition test and questionnaire.** The recognition test and questionnaire procedures matched those of Experiment 1. Unlike in Experiment 1, no participant accurately identified the purpose of the experiment.

### 4.2. Results

#### 4.2.1. Categorizing sounds

A  $2(\text{Sound Type: real, fake}) \times 2(\text{Sentence Congruence: match, mismatch})$  analysis of variance (ANOVA) on response times revealed no effect of Sound Type,  $F(1, 79) = 1.32, p > .05$ , or Sentence Congruence,  $F(1, 79) = 1.97, p > .05$ . As found in Experiment 1, Sound Type and Sentence Congruence interacted,  $F(1, 79) = 5.78, p < .05, \eta^2 = .02$ , as depicted in Fig. 1b. To directly compare the match and mismatch conditions, we conducted two paired t-tests, one in each of the two Sound Type conditions. Within the real condition, the match condition showed faster response times ( $M = 2.03$  s,  $SE = .08$ ) relative to the mismatch condition ( $M = 2.29$  s,  $SE = .11$ ),  $t(79) = 3.15, p < .01, d = .35$ ; this effect was not replicated in the fake condition,  $t(79) = .51, p > .05$ .

To further rule out the potential effects of task-related strategies, we conducted a  $2 \times 2 \times 3$  ANOVA with the addition of Segment (first, second and third) as the third repeated-measures variable. If readers are only mentally simulating sounds due to task demands, then the  $2 \times 2$  interaction should only occur in later (second and third)

segments. This was not the case, as demonstrated by the absence of a three-way interaction,  $F(2, 106) = .54, p > .05$ ; in fact, in all three segments, paired t-tests showed faster real sound categorization in the matching versus mismatching conditions (Block 1:  $t(70) = 2.22, p < .05, d = .26$ ; Block 2:  $t(74) = 2.05, p < .05, d = .23$ ; Block 3:  $t(70) = 2.12, p < .05, d = .24$ ).

#### 4.2.2. Recognition test

Overall, recognition accuracy was moderate ( $M = .70, SE = .02$ ) and false alarms were moderately low ( $M = .26, SE = .01$ ). Recognition accuracy did not vary as a function of our experimental conditions (all  $p$ 's  $> .26$ ), nor did recognition response times (all  $p$ 's  $> .11$ ).

### 4.3. Experiment 2 discussion

Our second experiment separated the reading and sound categorization tasks to further assess whether readers would perform auditory mental simulations without the potential influence of task demands. With this design, we replicated Experiment 1 results with faster real sound categorization when the sound matched rather than mismatched one of the sentences read in the immediately preceding block. Further, this effect occurred during the first block of experimental trials, suggesting that readers perform spontaneous auditory mental simulations prior to having the opportunity to develop task-related strategies. Overall, these effects also point to the lasting nature of auditory simulations formed during text comprehension. Relative to Experiment 1, there was a longer duration between reading and sound categorization; even with this gap, the auditory simulations persisted in memory and affected subsequent categorization response times (for similar results in the visual modality, see Wassenburg & Zwaan, in press).

## 5. General discussion

The present experiments assessed whether readers perform auditory mental simulations during sentence comprehension. In our first experiment, we found that readers correctly categorize real sounds faster when those sounds match a concept previously activated during reading. These results extend research demonstrating that visual or action-based sensorimotor activation during language comprehension can facilitate performance on subsequent verification tasks using congruent rather than incongruent stimulus properties (e.g., Glenberg & Kaschak, 2002; Glenberg et al., 2008; Stanfield & Zwaan, 2001; Yaxley & Zwaan, 2007), and extend these findings to the domain of auditory perception. If readers had not mentally simulated sounds during sentence comprehension, then sound categorization performance would not have varied as a

function of whether sounds matched or mismatched those implied by the sentences. In addition, our results are consistent with previous findings that readers use overlapping processing mechanisms for the representation of motion from auditory and language-based stimuli (Kaschak et al., 2006).

These results lend support to Perceptual Symbols Theory, which posits that conceptual representations are built from the activation of perceptual symbols that involve the full range of human sensory capabilities (Barsalou, 1999). That is, readers perform multimodal simulations that include the reinstatement of the very same neural patterns activated (i.e., experiential traces; Zwaan, 2004) during actual perception and action. Perceptual Symbols Theory thus predicts the spontaneous activation of auditory simulations that conceptually represent the act of hearing what is being described. We provide the first empirical evidence that readers activate specific auditory experiential traces, expanding the scope of mental simulations during reading and providing further evidence that amodal symbols frameworks cannot fully account for human performance. Indeed an amodal framework would predict that sound categorization performance would be unaffected by the content of processed language.

The present results also support the notion that modality-specific conceptual representations, in this case incorporating auditory imagery, can be activated through alternate modalities (Pecher, Zanolie, & Zeelenberg, 2007; van Dantzig et al., 2008). That is, the process of reading through the visual modality can promote the development of conceptual representations that incorporate different (and even non-dominant) modalities, evenuating in perceptually-rich situation models of what is being described in a text. This finding converges with behavioral and cognitive neuroscience evidence that comprehending concepts activates modality-specific brain areas (i.e., Kemmerer, Gonzalez Castillo, Talavage, Patterson, & Wiley, 2008; Martin, 2007; Pulvermuller, 2005, 2008; Tettamanti et al., 2005).

Our second primary finding relates to whether readers perform auditory simulations spontaneously and without the influence of task demands. Some have suggested that experimental tasks that include a meaningful overlap of language (e.g., “Through the fogged goggles, the skier could hardly identify the moose.”) and perception (e.g., a picture of a moose; Yaxley & Zwaan, 2007) may encourage the strategic use of mental imagery (i.e., Ditman et al., 2010; Machery, 2007; van Dantzig et al., 2008). Under this assumption, modal symbols only underlie language comprehension when readers are under experimental demands to form mental images of described events; without such demands, this view posits that amodal symbols underlie relatively naturalistic language comprehension. Whereas the present results cannot rule out the potential influence of task demands in previous studies, they do provide strong evidence that readers perform mental simulations spontaneously and in the absence of language and perception tasks that are simultaneous or highly related. Two primary results support this conclusion. First, Experiment 1 participants showed faster sound categorization when the sound was congruent rather than incongruent with a sound implied by a preceding sentence, even though the sound task was ostensibly unrelated and sequential (rather than simultaneous) to the reading task. Second, Experiment 2 participants showed this same effect when sound categorization was temporally separated from reading. Further evidence comes from the fact that Experiment 2 readers showed the match effect on the first experimental segment, prior to having an opportunity to develop strategies to facilitate task performance. The persistence of these results across the temporal lag and in the presence of multiple potential auditory simulations per experimental block, demonstrates the persistence and strength of auditory simulations over time.

Our results support a body of evidence demonstrating that language comprehension and sensorimotor activation are closely related, and that the mental simulation of described events is

multimodal and likely encompasses the entire range of human sensory capabilities (Barsalou, 2008; Zwaan, 2004). Current work in our laboratory is examining whether readers spontaneously activate tactile representations by examining whether such representations might drive judgments of texture properties. Importantly, though participants are able to develop motor and visual mental imagery under instruction (Jeannerod, 1994; Kosslyn, 1980), we add to a growing body of evidence demonstrating that readers generate detailed and multimodal mental simulations even in the absence of tasks promoting mental imagery.

## Acknowledgements

The opinions expressed herein are those of the authors and not necessarily of the United States Army. We thank Aaron Gardony and Vivien J. Lim for their assistance with stimuli development and data collection.

## References

- Alexander, J. E. D., & Nygaard, L. C. (2008). Reading voices and hearing text: Talker-specific auditory imagery in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 34(2), 446–459.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–609.
- Barsalou, L. W. (2005). Situated conceptualization. In H. Cohen, & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 619–650). St. Louis: Elsevier.
- Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, 59, 617–645.
- Bergen, B., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, 31, 733–764.
- Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, 32, 863–873.
- Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When you and I share perspectives: Pronouns modulate perspective-taking during narrative comprehension. *Psychological Science*, 20, 27–32.
- Burgess, C., & Lund, K. (1997). Modeling parsing constraints with high-dimensional context space. *Language and Cognitive Processes*, 12, 177–210.
- Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 687–696.
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240–247.
- Ditman, T., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition*, 115, 172–178.
- Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2008). Time travel through language: Temporal shifts rapidly decrease information accessibility during reading. *Psychonomic Bulletin & Review*, 15, 750–756.
- Fiedler, K. (2009). On embodied cognition and mental simulation: A meta-theoretical comment to Zwaan's treatise. *European Journal of Social Psychology*, 39, 1156–1159.
- Fincher-Kiefer, R. (2001). Perceptual components of situation models. *Memory & Cognition*, 29, 336–343.
- Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, 61, 825–850.
- Fodor, J. A. (1975). *The Language of Thought*. New York: Crowell.
- Fodor, J. A. (2000). *The Mind Doesn't Work that Way*. Cambridge, MA: MIT Press.
- Gernsbacher, M. A. (1990). *Language Comprehension as Structure Building*. Hillsdale, NJ: Erlbaum.
- Gernsbacher, M. A., & Foertsch, J. (1999). Three models of discourse comprehension. In S. Garrod, & M. J. Pickering (Eds.), *Human language processing* (pp. 283–299). East Sussex, UK: Psychology Press.
- Glenberg, A. M. (1997). What memory is for. *Behavioral & Brain Sciences*, 20, 1–55.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558–565.
- Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology*, 61, 905–919.
- Glover, S., Rosenbaum, D. A., Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research*, 154, 103–108.
- Graesser, A. C., Gernsbacher, M. A., & Goldman, S. (Eds.). (2003). *Handbook of discourse processes*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Horton, W. S., & Rapp, D. N. (2003). Out of sight, out of mind: Occlusion and the accessibility of information in narrative comprehension. *Psychonomic Bulletin & Review*, 10, 104–110.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17, 187–245.



- Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, 43, 508–529.
- Kaschak, M. P., Madden, C. J., Theriault, D. J., Yaxley, R. H., Aveyard, M. E., Blanchard, A. A., et al. (2005). Perception of motion affects language processing. *Cognition*, 94, B79–B89.
- Kaschak, M. P., Zwaan, R. A., Aveyard, M., & Yaxley, R. H. (2006). Perception of auditory motion affects language processing. *Cognitive Science*, 30, 733–744.
- Kemmerer, D., Gonzalez Castillo, J., Talavage, T., Patterson, S., & Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: Evidence from fMRI. *Brain and Language*, 107, 16–43.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163–182.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: MIT Press.
- Kurby, C. A., Magliano, J. P., & Rapp, D. N. (2009). Those voices in your head: Activation of auditory images during reading. *Cognition*, 112, 457–461.
- Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. Chicago and London: University of Chicago Press.
- Landauer, T. K. (2002). On the computational basis of learning and cognition: Arguments from LSA. In N. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 43–84). : Academic Press.
- Landauer, T. K., Laham, D., & Foltz, P. W. (1998). Learning human-like knowledge by singular value decomposition: A progress report. In M. I. Jordan, M. J. Kearns, & S. A. Solla (Eds.), *Advances in Neural Information Processing Systems* (pp. 45–51). Cambridge, MA: MIT Press.
- Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2007). *Handbook of Latent Semantic Analysis*. Mahwah, NJ: Erlbaum.
- Machery, E. (2007). Concept empiricism: A methodological critique. *Cognition*, 104, 19–46.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, 11, 194–201.
- Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language*, 26, 165–187.
- Pecher, D., Zanolie, K., & Zeelenberg, R. (2007). Verifying visual properties in sentence verification facilitates picture recognition memory. *Experimental Psychology*, 54, 173–179.
- Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, 83, 157–171.
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6(7), 576–582.
- Pulvermüller, F. (2008). Grounding language in the brain. In M. de Vega, A. Graesser, & A. M. Glenberg (Eds.), *Symbols, embodiment, and meaning* (pp. 85–116). Oxford: Oxford University Press.
- Pylyshyn, Z. W. (1984). *Computation and cognition*. Cambridge MA: MIT Press.
- Pylyshyn, Z. W. (2002). Mental Imagery: In Search of a Theory. *Behavioral and Brain Sciences*, 25, 157–182.
- Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with “The language-as-fixed-effect fallacy”: Common misconceptions and alternative solutions. *Journal of Memory and Language*, 41, 416–426.
- Rayner, K. (1998). Eye movements in reading and information processing: Twenty years of research. *Psychological Bulletin*, 124, 372–422.
- Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science*, 27, 767–780.
- Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, 4, 546–550.
- Spivey, M. J., Tyler, M., Richardson, D. C., & Young, E. (2000). Eye movements during comprehension of spoken scene descriptions. *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society* (pp. 867–872). Mahwah, NJ: Erlbaum.
- Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153–156.
- Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17, 273–281.
- Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica*, 116, 185–203.
- Van Dantzig, S., Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2008). Perceptual processing affects conceptual processing. *Cognitive Science*, 32, 579–590.
- Wassenburg, S.I., & Zwaan, R.A. (in press). Readers routinely represent implied object rotation: The role of visual experience. *Quarterly Journal of Experimental Psychology*.
- Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition*, 105, 229–236.
- Zwaan, R. A. (1999). Situation models: The mental leap into imagined worlds. *Current Directions in Psychological Science*, 8, 15–18.
- Zwaan, R. A. (2004). He immersed experienter: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *Psychology of Learning and Motivation*, Vol. 44 (pp. 35–62). San Diego, CA: Academic Press.
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6, 292–297.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13, 168–171.
- Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1–11.
- Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic-relatedness judgments. *Psychonomic Bulletin & Review*, 10, 954–958.