

The concurrent use of three implicit measures (eye movements, pupillometry, and event-related potentials) to assess receptive vocabulary knowledge in normal adults

Kerry Ledoux · Emily Coderre · Laura Bosley ·
Esteban Buz · Ishanti Gangopadhyay · Barry Gordon

© Psychonomic Society, Inc. 2015

Abstract Recent years have seen the advent and proliferation of the use of implicit techniques to study learning and cognition. One such application is the use of event-related potentials (ERPs) to assess receptive vocabulary knowledge. Other implicit assessment techniques that may be well-suited to other testing situations or to use with varied participant groups have not been used as widely to study receptive vocabulary knowledge. We sought to develop additional implicit techniques to study receptive vocabulary knowledge that could augment the knowledge gained from the use of the ERP technique. Specifically, we used a simple forced-choice paradigm to assess receptive vocabulary knowledge in normal adult participants using eye movement monitoring (EM) and pupillometry. In the same group of participants, we also used an N400 semantic incongruity ERP paradigm to assess their knowledge of two groups of words: those expected to be known to the participants (high-frequency, familiar words) and those expected to be unknown (low-frequency, unfamiliar words). All three measures showed reliable differences between the known and unknown words. EM and pupillometry thus may provide insight into receptive vocabulary knowledge similar to that from ERPs. The development of additional implicit assessment techniques may increase the feasibility of receptive vocabulary testing across a wider range of

participant groups and testing situations, and may make the conduct of such testing more accessible to a wider range of researchers, clinicians, and educators.

Keywords Eye movements · Pupillometry · Event-related potentials · Receptive vocabulary

One of the great challenges in the study of cognition and learning is to know what an individual knows. What would seem to be the most direct method—just asking them—is fraught with many limitations. For example, the representations in the cognitive architecture and the processes that operate on them are not always available to conscious access. Even to the degree that they are, it may be difficult for the typical adult participant to describe them using commonplace language. And asking is simply impossible for those with limited or absent verbal communication abilities, such as infants, small children, those with some kinds of developmental disabilities, and nonhuman animals.

For this reason, other methods to assess learning and cognition have been developed that rely instead on observations of the participants' behavior. Inferences are then drawn between these behavioral measures and the more elusive constructs of interest. One such behavioral measure is reaction time (RT): Given the assumption that cognitive processes unfold in time, the measurement of how long it takes an individual to respond to stimuli that vary along different dimensions can provide insight into the number of processes being engaged, the difficulty of those processes, or the time needed to access the stored representations, all of which might meaningfully differ across experimental conditions (Donders, 1969; Posner, 2005). Another example is the habituation paradigm, used to study cognition in infants, in which looking time is used as a measure of interest or stimulus novelty in babies: Babies will visually attend to a stimulus until they no

K. Ledoux · E. Coderre · L. Bosley · E. Buz · I. Gangopadhyay ·
B. Gordon
Cognitive Neurology/Neuropsychology Division,
Department of Neurology, Johns Hopkins University
School of Medicine, Baltimore, MD, USA

B. Gordon
Department of Cognitive Science, Johns Hopkins University,
Baltimore, MD, USA

K. Ledoux (✉)
Cognitive Neurology/Neuropsychology, Johns Hopkins University,
1629 Thames St. Ste. 350, Baltimore, MD 21231, USA
e-mail: kledoux1@jhmi.edu

longer perceive it as novel, and they will look away when they tire of it. Making small alterations to a stimulus after the child has looked away from it allows researchers to determine, on re-presentation, whether the child is aware of the changes (if he or she spends time looking at it again) or not (if he or she doesn't; Colombo & Mitchell, 2009; Fantz, 1964). These and other behavioral methods have an extensive history of use in the fields of cognitive psychology, cognitive neuroscience, and education, and have contributed greatly to our current understanding of cognition and learning.

These methods are not without limitations themselves, however. One important limitation is that behavioral techniques often require an understanding of task instructions and/or the execution of complex behaviors in responding, making them difficult or impossible to use with certain participant populations. Another limitation is the ability to generalize their use to participant populations other than normal adults. For example, making inferences about age-related changes in cognitive processing from RT studies can be difficult, because such changes may be confounded with age-related changes in motor responses. Even the habituation technique described above, which has been used successfully to study cognition in infants, may present difficulties of interpretation for groups in whom looking behavior may be unreliable, such as low-functioning individuals with autism. Finally, many of these behavioral techniques depend on a participant's motivation to engage in and complete the task, something that again might vary tremendously across participant groups (and even across testing sessions, for individual participants).

For this reason, recent years have seen an emphasis on the development of assessment techniques that do not necessarily rely on an explicit behavioral response. These more implicit assessment methods (which include techniques such as functional magnetic resonance imaging, event-related potentials, eye movement monitoring, and pupillometry, among others) have the advantage of being useable even with individuals in whom more overt verbal or behavioral responses would be difficult or impossible to reliably obtain. Thus, they may prove especially useful in the study of cognition and learning in infants, small children, and patient populations, groups that frequently have been underrepresented in studies of cognition due to the difficulty of testing them.

Event-related potentials in the study of receptive vocabulary knowledge

An example of the development of an implicit technique to study an area of learning comes from the prolific recent use of event-related potentials (ERPs) to study receptive vocabulary knowledge. ERPs are event-locked changes in the scalp-recorded electroencephalogram (EEG). ERPs provide information with very fine temporal resolution (on the order of

milliseconds) about how cognitive processes unfold in real time. Most importantly, ERPs can be meaningfully observed even in the absence of an overt behavioral response.

Separable ERP components have been reliably associated with separable aspects of cognitive processing; the evaluation of different ERP components can give a direct indication of the involvement of individual cognitive functions. One such component that has been reliably associated with cognitive operations is the N400, a negative-going deflection that peaks approximately 400 ms after the onset of a meaningful stimulus (such as a word or picture; Kutas & Hillyard, 1980). The N400 has been demonstrated to index semantic integration processing, by which the meaning of new stimuli is understood as being a part of, and integrated with, the current semantic context. Stimuli that are easier to integrate with their preceding context (for example, those that are semantically congruent with their context) elicit a smaller-amplitude N400 than words or pictures that are more difficult to integrate (for example, those that are semantically incongruent) with their context (Brown & Hagoort, 1993; Holcomb, 1993; Kutas & Federmeier, 2000; Van Petten & Kutas, 1991). The difference in the amplitude of the N400 between congruent and incongruent conditions has been called the N400 congruity effect.

This elicitation of the N400 congruity effect has been exploited to study receptive vocabulary knowledge by pairing a word with the meaningful context of a concurrently presented picture. A reduction of the amplitude of the N400 component is observed when the picture matches the named word (reflecting the ease of integrating the matching stimuli) relative to when it does not (reflecting the greater difficulty in integrating the incongruent stimuli). Critically, the integration of the auditory word with the picture context, and the resultant reduction of the N400 in cases of congruity, depends upon the listener/viewer having knowledge of the word's meaning. To the extent that the word is unknown to the participant, the integration between the word and its context cannot be eased by congruity because semantic knowledge cannot be brought to bear on the situation. In this way, an N400 congruity effect is predicted between spoken words and pictures, but *only* for words that are known to the participant. In the case of unknown words, no reduction of the amplitude of the N400 component would be expected, because the participant cannot use semantic knowledge to ease integration in the congruent condition. In other words, there should not be a difference in the ERPs for the congruent and incongruent conditions for unknown words because if the participant does not know the meaning of the word, he or she cannot assess whether the word and picture match.

A number of studies have confirmed these predictions. For example, Connolly and colleagues (Byrne et al., 1999; Connolly & D'Arcy, 2000; Connolly, D'Arcy, Newman, & Kemps, 2000; D'Arcy et al., 2003; Marchand, D'Arcy, & Connolly, 2002) used this type of N400 congruity paradigm

to assess receptive vocabulary in a series of studies with various participant groups, including healthy adults, typically developing children, adults with aphasia, and a child with cerebral palsy (for whom motor activity, and thus behavioral response, was limited). They used a variety of behavioral measures to estimate each participant's receptive vocabulary level, then presented congruent and incongruent pictures with words that were within or beyond that participant's vocabulary level. In each case, a larger N400 was observed to the auditory word in the incongruent condition than in the congruent condition—but only for words that were within the participant's vocabulary level. Other research has revealed similar results with a variety of participant groups (see, for example, Friedrich & Friederici, 2004, 2005a, b, 2010; Henderson, Baseler, Clarke, Watson, & Snowling, 2011; Torkildsen et al., 2008), and has even demonstrated the elicitation of the effect following training on new words (Friedrich & Friederici, 2008; Junge, Cutler, & Hagoort, 2012; Key, Molfese, & Ratajczak, 2006; Ojima, Nakamura, Matsuba-Kurita, Hoshino, & Hagiwara, 2010; Torkildsen et al., 2009). These findings thus support the utility of ERP measures to help discriminate sets of known words from sets of unknown words, and demonstrate the capability of this technique to be used in the testing of a wide variety of participant groups (including those who may otherwise have struggled to make overt behavioral responses).

Despite this demonstrated utility, the ERP paradigm carries some important potential limitations to the study of receptive vocabulary knowledge. First, this technique does not easily allow examination of the brain's response to a single item. Because the signal (the electrical activity of the brain) is relatively weak when compared to the multiple sources of noise in the EEG recording (such as eye movements, blinks, and muscle activity, all of which contribute to the recorded electrical signal), it is only through averaging the time-locked signal to multiple trials of a like type (or from the same experimental condition) that the brain potential can be sufficiently isolated from the noise. The greater the number of trials included in the average, the better the chance of eliminating more of the noise and observing more of the brain's activity. For this reason, studies that have used ERP paradigms to study receptive vocabulary have done so by comparing the averaged response across many trials of like words. When this is done, we see that the ERPs to known words differ from those to unknown words. This is very useful information, to the extent that researchers and experimenters can determine a priori pools of known and unknown words. Yet imagine the case of a clinician or a teacher, who wishes to use some objective measure to determine precisely which words are known and which are not, on a single word basis. This kind of determination from ERP data is not possible, because the observed signal to a single word on a single trial is

simply too noisy. (One possible solution to this problem would be to run multiple trials of a single word and to average these together, although this approach also contains potential limitations, such as the fact that ERPs are also very sensitive to lexical repetition.) For this reason, the development of other implicit measurement techniques that provide benefits similar to those of ERPs, but that might also allow the examination of responses to single words, would be useful.

A second potential limitation to the use of an ERP paradigm to assess vocabulary is that ERP equipment (and the training necessary to learn to use it) is not necessarily readily accessible to the wide variety of professionals (clinicians, teachers, speech-language pathologists, etc.) and to the families who work with them, who might benefit from having a greater knowledge of an individual's receptive vocabulary level. ERP equipment has certainly become more affordable in recent years (even for very high-density electrode systems), and because of this and an increase in the application of ERPs to the study of cognition, more and more research groups in psychology, speech-language science, and other areas of cognitive neuroscience do have the potential to use this technology to study cognitive function. However, the cost remains relatively prohibitive for widespread access and use, especially compared to other technologies and methods of investigation.

Finally, some participant groups (such as infants, young children, and patients with acquired or developmental disorders) may prove less amenable to ERP testing, which may take longer than other methodologies (due to the need to acquire many trials) and which often involves the lengthy application of equipment that may prove intolerable to some (such as individuals with autism, who are frequently observed to dislike things being placed on their heads). Additionally, for reasons mentioned above, the isolation of the brain's electrical activity generally benefits from a reduction in electrical activity from other sources, such as muscles or the eyes. Keeping such sources of extraneous electrical noise (artifacts) out of the ERP is usually best accomplished through instructions to participants, for instance, to refrain from moving the body and the eyes and to refrain from blinking during critical portions of the trial. Participant groups who have more difficulty understanding such instruction or complying with them will necessarily have noisier data. Traditionally, the detection of artifacts during critical trials resulted in data loss, as such trials would be rejected from the final analysis. Recently, data-analytic procedures have been developed that allow for artifact correction (in place of artifact rejection), but even these methods cannot guarantee the removal of extraneous noise and often still result in data loss that may make the inclusion of participants from certain groups difficult or impossible.

Other implicit assessment techniques: Eye movement monitoring and pupillometry

For these reasons, in the present study we aimed to determine whether the assessment of receptive vocabulary knowledge could be extended to other implicit measurement techniques—specifically, eye movement monitoring (EM) and pupillometry. Both of these techniques have been used in the study of a variety of cognitive processes, as we discuss below. Additionally, they both may prove capable of providing more stable or reliable information about single words from single trials, which could be of tremendous benefit in determining which individual words are known to an individual participant. Furthermore, the equipment used to collect EM and pupillometry data is generally available at lower costs than ERP equipment, and may be available to some clinicians or instructors to whom ERP equipment is not. To the extent that EM and pupillometry data could be collected from paradigms that complement those used to study receptive vocabulary using ERPs, this might extend the application of implicit techniques generally to wider participant populations.

Eye movement monitoring EM and pupillometry both have long histories of application to the study of various aspects of cognitive processing. Eye movements have long been taken to reflect current cognitive operations (Just & Carpenter, 1976; Rayner, 1998), and thus have been used extensively to study various aspects of cognition, especially language. Although much of this application has been in the study of reading, eye movements have also played an important role in our understanding of the interface between spoken word recognition and aspects of language processing such as phonology or semantics. For example, Cooper (1974) demonstrated that participants would move their eyes to various pictures of objects as they heard those objects named in a story. More recently, similar demonstrations have come using the visual world paradigm (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Tanenhaus, 2007; Tanenhaus, Magnuson, Dahan, & Chambers, 2000; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In this paradigm, participants are asked to look at visual displays while listening to speech (which might include explicit instructions to manipulate the objects, either with their hands or with a computer mouse, or might be presented in a more passive way without instruction). These studies have consistently shown a tight time-locking between the unfolding of the auditory signal and participants' eye movements, such that participants will move their eyes to named objects in the display as soon as they can recognize even a minute fragment of the auditory stimulus. These types of paradigms have been used to study the time course of spoken word recognition, for example, by identifying the time at which two objects with names that

share speech onsets become disambiguated, both in adults (Allopenna, Magnuson, & Tanenhaus, 1998; Dahan, Magnuson, Tanenhaus, & Hogan, 2001) and in children (Fernald, Perfors, & Marchman, 2006; Fernald, Swingley, & Pinto, 2001; Swingley & Fernald, 2002; Trueswell, Sekerina, Hill, & Logrip, 1999). Recently, and importantly for our purposes, such paradigms have also been used in the study of semantic activation during speech perception (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2005, 2011; Myung, Blumstein, & Sedivy, 2006; Myung et al., 2010; Yee, Huffstetler, & Thompson-Schill, 2011). For example, Yee and Sedivy (2006) showed that as participants heard a spoken word, their eyes were more likely to move not only to the named object in the display, but also to a semantically related object (for instance, if the spoken word was *lock*, eye movements were more likely to a picture of a lock and also to a picture of a key, relative to unrelated pictures in the display).

Odekar, Hallowell, Kruse, Moates, and Lee (2009) used a similar visual-world-type paradigm to determine whether patterns of eye movements could be used as valid and reliable indicators of semantic priming. Participants were presented with a printed prime word (such as *marriage*), followed by a display of three objects. On related trials, one of the three objects (e.g., a ring) shared a semantic/associative relationship with the prime, whereas the other two (for example, a nail and an ear) did not. Odekar and colleagues found that participants were faster to look at a given picture when presented in the related condition (i.e., preceded by a related prime word) as compared to when the same picture appeared in the unrelated condition (i.e., preceded by an unrelated prime word). Specifically, participants looked longer at related pictures both on average (mean fixation duration) and on initial processing measures (first fixation duration). They were also faster to move their eyes to a picture for the first time (latency to first fixation) in the related condition. Additionally, proportional measures, such as proportion of fixation durations on the target picture, showed that participants spent relatively greater amounts of total looking time over the course of the trial on pictures that were related to the preceding word. Finally, they found that a higher percentage of first fixations across trials were made to the named object on related trials, relative to unrelated trials. These results suggest that semantic priming is indeed reflected in differential patterns of eye movements in a visual-world-type paradigm. On the basis of these results, we hypothesized that the semantic relationship between an auditory word and its visual image would similarly affect eye movements, but only to the extent that that relationship was known to the participant. For both semantically related items and for known words, we expect to see an advantage in processing that is conferred by the knowledge of a semantic match between the auditory cue and the picture. Such

an advantage would not be conferred in the cases in which the items do not share a semantic relationship (either because they do are not close in the semantic network, or because one does not have the necessary knowledge to determine if they are related), and this should lead to greater relative difficulties in processing. In this way, we expected to see similar differences as those observed by Odekar and colleagues when comparing eye movements in our known and unknown word conditions.

Pupillometry Pupillometry measures changes in pupil dilation to study various aspects of cognitive processing. Pupillary dilation can be affected by many environmental or participant-internal events (such as changes in lighting, emotional arousal, or the onset of stress; Beatty & Lucero-Wagoner, 2000; Goldwater, 1972; Hess, 1965; Hess, Seltzer, & Shlien, 1965; Loewenfeld, 1993). Importantly, however, changes in pupil size are also observed in relation to the demands elicited by cognitive tasks, and these changes have been shown to occur independently of other influences (Goldwater, 1972; Karatekin, Marcus, & Couperus, 2007). Such changes are generally observed by time-locking changes in pupil diameter to the onset of stimuli that elicit various cognitive processes, and thus are often referred to as task-evoked pupillary reflexes (Beatty, 1982; Kahneman & Beatty, 1966; Kahneman, Beatty, & Pollack, 1967). Such task-specific changes in pupillary diameter have long been associated with attentional engagement and information processing: pupillary dilation has been shown to increase with task difficulty in many paradigms, and has thus been taken as a measure of resource recruitment (Beatty, 1982; Hess & Polt, 1964; Hoeks & Levelt, 1993; Kahneman & Beatty, 1966). Recently, Kuipers and Thierry (2011, 2013) examined pupillary responses recorded during an N400 picture–word congruity paradigm with high frequency words. For adults, congruent picture–word pairings elicited both a reduction in the amplitude of the N400 and smaller pupil sizes, relative to incongruent pairings. In a second study comparing monolingual and bilingual children, similar N400 congruity effects were observed in both groups. However, only the bilingual children showed the congruity effect in the pupillometry measures, suggesting that there may be developmental changes in resource recruitment during this task that might emerge earlier in children who speak more than one language.

For our purposes, we measured pupillary dilation during a four-alternative forced-choice task. We expected that participants would show greater cognitive resource recruitment when asked to select the picture that matched an unknown auditory word, relative to the very well-known and overlearned pairings between the known words and their visual depictions used in our study. We therefore anticipated that we would see larger changes in pupillary dilation in the unknown condition than in the known condition.

The present study

In the present study, we used all three of these implicit assessment techniques—EM, pupillometry, and ERP—to assess receptive vocabulary in a group of normal adult participants. The use of EM to study receptive vocabulary was novel to this study; two previous studies have used pupillometry for this purpose, but in a different paradigm than we employed here. We presented participants with two tasks involving pictures and auditory tokens of two sets of words: words that we expected would be known to all of the participants (such as *circle* and *dog*) and words that were expected to be unknown to most of the participants (such as *bilby* and *loquat*). In the *forced-choice task*, participants saw four pictures on the screen and heard one named; they were asked to use the mouse to select the named picture while EM and pupillary diameter data were recorded. In the *congruity task*, participants saw one picture on the screen and heard an auditory token that either matched (congruent condition) or did not match (incongruent condition) the picture. Participants indicated by button-press whether the picture matched or did not match the spoken word. ERP data were acquired during this task. In both tasks, half of the trials involved “known” words and half involved “unknown” words.

We made specific predictions for each of the three implicit measures. For EM, we predicted that eye movements would be faster and more reliable to the named picture for known words; for unknown words, looking behavior was expected to be more random and variable. Specifically, on the basis of the previous study conducted by Odekar and colleagues (2009) for semantic priming, we expected that measures of looking time (mean fixation duration, first fixation duration, and first dwell duration) would be longer to known items than unknown items. We also expected that participants would be faster to look at the named picture for the first time (latency to first fixation) and to return to the named picture once having moved the eyes away from it (latency to first refixation) for known items relative to unknown items. We expected that the proportional measures of time spent fixating the picture (proportion of fixation duration on stimulus) and of time spent looking at the named picture whether fixating or not (proportion of dwell time on stimulus) would both be greater for known words than for unknown words. Finally, we expected that the percentage of times that the named object would be the first and the last object fixated during a trial would be greater for known objects than for unknown objects. For pupillometry, we predicted that changes in pupillary dilation from baseline would be greater in the unknown condition, relative to the known, reflecting greater resource recruitment when the word’s meaning was unknown. For ERPs, we expected to replicate previous studies that have used the N400 to assess receptive vocabulary knowledge by showing an N400 congruity effect for known words; that is, the amplitude of the

N400 component of the ERP was expected to be larger in the incongruent condition, relative to the congruent condition, but only for words that were known to the participant. For unknown words, prior knowledge could not be used to determine the congruity between the auditory word and the picture, and therefore the amplitude of the N400 was expected to be approximately the same for the congruent and incongruent conditions. Thus, no N400 congruity effect was predicted for the unknown condition.

Method

Participants

The participants were 23 adult, right-handed native speakers of English between the ages of 19 and 61 ($M = 35$ years; 70 % male, 30 % female). All had self-reported normal or corrected-to-normal vision and hearing. None of the participants reported cognitive, learning, or neurological impairment, and none were currently taking medication that might affect neurological or cognitive functioning. Participants were recruited from the Johns Hopkins University and surrounding community. The experimental procedures had been approved by the Johns Hopkins School of Medicine Institutional Review Board, and all participants gave written informed consent before participating in the experiment. All were monetarily compensated for participating.

Materials

Participants completed two separate tasks (a forced-choice task and a congruity task; see below) using the same set of 160 words and pictures. Eighty of the words (hereafter called “known words”) were very high frequency (as assessed by the Subtlex US database; Brysbaert & New, 2009; $M = 3.14$, $SD = 0.6$), and were expected to be familiar to even very young children; these included words such as *cat*, *airplane*, and *camera*. The remaining 80 words (“unknown words”) were low in frequency ($M = 0.85$, $SD = 0.5$), relatively unfamiliar words that were not expected to be known by the majority of the adult participants. Pretesting of these materials with a separate group of normal adults from the Johns Hopkins University community ($n = 15$) confirmed that these words were relatively unknown to this group. Examples of words in this set included *cherimoya*, *agouti*, and *cainito*. All words were highly imageable. High-resolution, color digital pictures were selected to represent each word. Pretesting with a separate group of normal adult participants ($n = 3$) ensured the suitability of the images as representations of their corresponding concepts. High-quality, digital auditory recordings of a female speaker pronouncing the name of each of the items were made using Audacity software, and were edited using Computerized

Speech Lab Model 4150 software (KayPENTAX). The edited tokens ranged in length from 500 to 1,200 ms. The recordings were transcribed by two speech-language pathology graduate students naïve to the purposes of the experiment. Their transcriptions were compared to that of a speech-language pathologist who assisted with the recordings to ensure that the words were completely audible. Volume was normalized across the tokens.

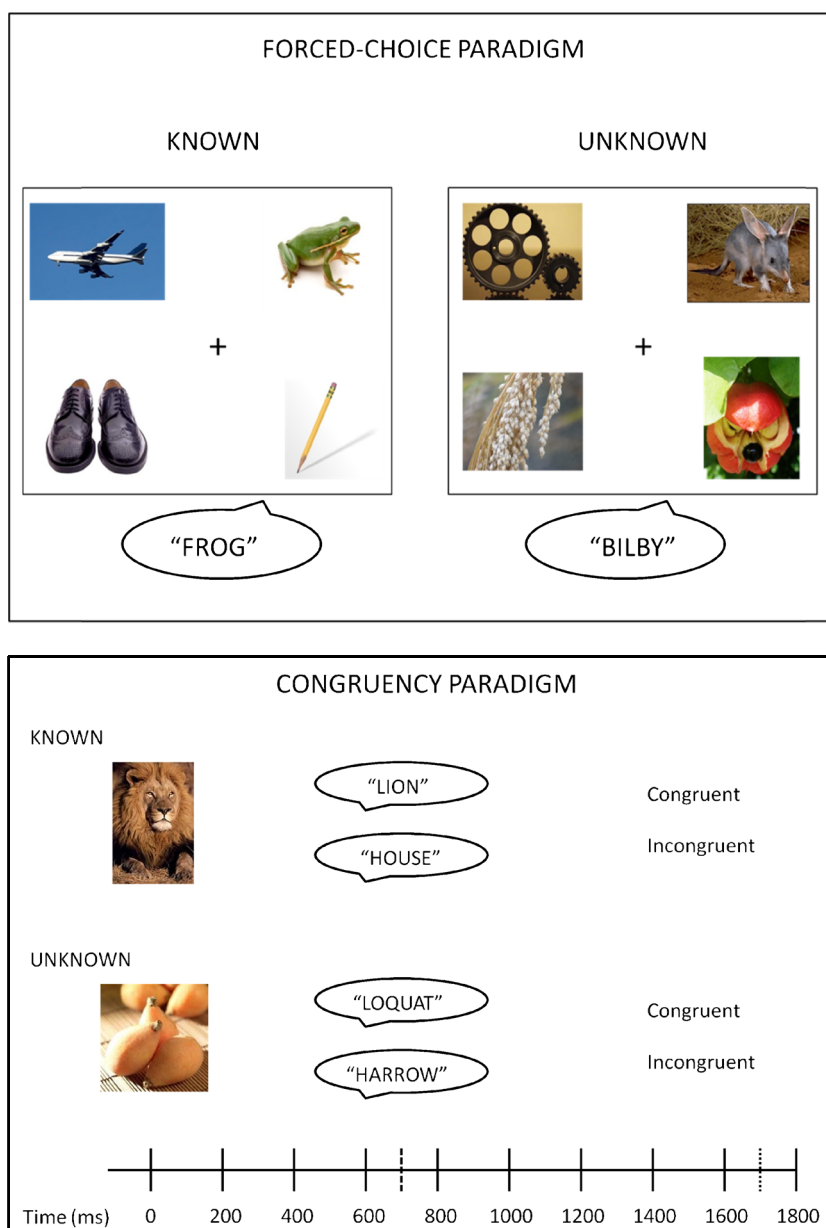
Procedure

Participants completed two testing sessions. In one, participants completed either the forced-choice task (during which EM and pupillometry data were collected) or the congruity task (during which ERPs were recorded), along with the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007). In the second session, participants completed the second of the two tasks (forced choice or congruity), along with a word familiarity rating task. We chose to collect EM and PD data during the forced-choice task, and ERP data separately during the congruity task, to maintain consistency with previous studies that have used similar paradigms to examine semantic processing (as described in the introduction). The three tasks are described in further detail below.

Forced-choice task In the forced-choice recognition task, presented in E-Prime (version 2.0.8.74), participants were asked to use the mouse to select one of four pictures presented simultaneously on a computer screen after having heard one of the objects named. On each trial, participants saw a fixation cross at the center of the screen for 1,000 ms. The fixation cross remained on the screen as the four pictures appeared, one in each corner of the screen (see Fig. 1). Twenty milliseconds later, participants heard one of the pictures named. On all trials, the three distractor items were drawn from the same knowledge category (known/unknown) to prevent participants from using a process of elimination on unknown trials. The pictures remained on the screen until the participant selected one with a mouse click, or for a maximum of 5 s. There were 160 trials, one per experimental item. These were presented in eight blocks of 20 trials, in which half were known targets and half were unknown (pseudorandomized within blocks). One practice trial served to familiarize participants with the paradigm at the start of the experiment. We simultaneously collected eye movement and pupillometry data using an ASL Model 504 eyetracking system. The pupil diameter was measured horizontally and was recorded every 17 ms in pixels. Reaction times and accuracy were also recorded.

Congruity task In the congruity paradigm, also presented in E-Prime, a picture was presented on the computer screen, followed immediately by the auditory presentation of a single word, which either matched (congruent condition) or did not

Fig. 1 Schematic representations of the forced-choice paradigm (top) and the congruency paradigm (bottom). See the text for details of each. For the unknown condition in the forced-choice example shown above, the pictured objects are (clockwise, from top left) *pinion*, *bilby*, *ackee*, and *millet*. For the unknown condition in the congruency example above, the pictured items are *loquats*



match (incongruent condition) the pictured item (see Fig. 1). As in the forced-choice task, the mismatching pictures for the incongruent condition were chosen from the same knowledge category (known/unknown) as the auditory token, to avoid strategic responses based on a process of elimination. Additionally, care was taken to ensure that the incongruent word–picture pairs did not share the same initial phoneme. A red fixation point (presented for 1,000 ms) began each trial, followed by the presentation of the picture for 700 ms. The auditory token was then presented (varying in length from 500 to 1,200 ms). The picture remained on the screen throughout the duration of the auditory token and for another 1 s after its offset, during which time responses were prohibited. A response screen (indicated by a green fixation point) was then

presented for up to 5 s or until participants made a response. Participants used two buttons on a button box to indicate whether the auditory word and the picture matched or did not match. They were also instructed to keep their eyes fixated on the center of the screen, to move as little as possible, and to refrain from blinking during the presentation of the picture and the auditory token. This was done to minimize artifacts in the EEG signal. There were 320 trials, two per experimental item (one congruent pairing, one incongruent pairing). These were presented in eight blocks of 40 trials each, in which ten trials of each type (known–congruent, known–incongruent, unknown–congruent, and unknown–incongruent) were pseudorandomized. In an initial trial block, ten nonexperimental items were presented to familiarize participants with the

task. High-density ERPs were recorded during the congruity task at 250 Hz using a Geodesics 256-channel sensor net (see Fig. 2) and NetStation version 4.3. Impedances were kept under 50 k Ω , where possible.

Word familiarity task Participants were asked to participate in a word familiarity posttest after the EM, pupillometry, and ERP testing had been completed. In the posttest, the participants were presented with each of the 160 auditory tokens and asked to rate their familiarity with the word on a scale from 1 (*very unfamiliar*) to 9 (*extremely familiar*), with an additional option of 0 (*no familiarity whatsoever*).

Data processing and analysis

The data from each of the three implicit measures were processed and analyzed separately. For all analyses, effect sizes are reported as Cohen's d (for t tests) or eta-squared (η^2 , for analyses of variance [ANOVAs]), calculated for the within-subjects measures.

Eye movements Fixation data were analyzed using ASL Results (Applied Science Laboratories, 2009). Each presentation slide was divided into five areas of interest: the four picture stimuli and the fixation cross in the middle of the screen. After discarding any trials in which more than 50 % of the trial was not detected, an average of 78 % of unknown and 75 % of known trials remained for statistical analysis.

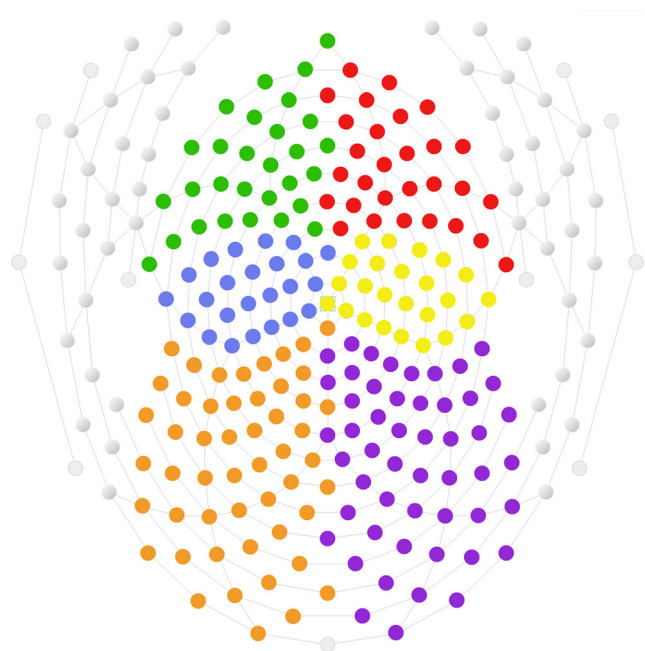


Fig. 2 The 256-channel electrode montage. Color is used to indicate the six electrode groupings used in the analysis: Green and red indicate frontal electrodes on the left and right, respectively; blue and yellow indicate central electrodes; and orange and purple indicate parietal electrodes

For each trial, we calculated a number of dependent variables derived from fixation and dwell time measures. Fixations were defined as periods of time (in milliseconds) for which eye gaze remained at a specific location on the screen. Fixation onsets were defined as a stable gaze duration of at least 100 ms and a visual angle variation of less than 1 degree. Fixation offsets were defined as three or more sequential samples that deviated from the fixation start location by more than 1 deg of visual angle. Dwell time was defined as the amount of time (in milliseconds) spent looking at the named picture, with or without fixation (i.e., the time that the eyes were in the region of interest of the named picture, whether they stayed in one spot long enough to reach the threshold for fixation). From these measures, the following dependent variables were derived:

Total number of fixations: the number of fixations made during the entire trial

Mean fixation duration: the average length of all fixations within the area of the named picture

First fixation duration: the length of the first fixation within the area of interest for the named picture

First dwell on stimulus: total time spent looking at the named picture during the first dwell

Latency to first fixation: the amount of time that passed before the first fixation on the named picture

Latency to first refixation: the amount of time that passed before a refixation occurred in the area of interest of the named picture (i.e., the amount of time to come back to fixate on the named picture after the eyes had left the region of this picture)

Proportion of fixation duration on the stimulus: the proportion of fixation duration time on the named picture relative to total fixation duration time (i.e., fixation duration on the named stimulus/total fixation durations for all four pictures)

Proportion of dwell time on stimulus: the proportion of the trial that was spent looking at the named picture, with or without fixation (i.e., dwell time on the named stimulus/length of the trial)

Percentage of trials first fixated: the percentage of trials (out of all good trials) on which the named object was the first picture to be fixated

Percentage of trials last fixated: the percentage of trials (out of all good trials) on which the named object was the last picture to be fixated

Pupillometry To convert the pixel measurement to millimeters, a scaling factor was calculated using a model eye provided by ASL to simulate the image received from a real eye. When viewed by the eye tracker optics, the model eye simulates a pupil image and corneal reflection. To calibrate the

pupil diameter, we positioned the model eye so that the white circle was at a normal eye distance from the optics, oriented so that the corneal reflection appeared below the pupil.

We used the Interface software to discriminate on the model pupil image, recorded the pupil diameter value on the computer screen digital display window, and computed the scale factor. To compute a scale factor (millimeters per eyetracker unit), the diameter of the white circle (4 mm) was divided by this value. To perform the analyses, the recorded pupil diameter values were converted to millimeters by applying this scale factor (value in millimeters = scale factor \times recorded value).

Pupillary responses were analyzed using Microsoft Office Excel 2007 and IBM SPSS Statistics 19. Prior to the statistical analyses, the data were cleaned by removing artifacts due to excessive blinking and by replacing small blinks by linear interpolation. Trials in which 20 or more data points in a row (340 ms or more) were missing due to a lack of fixations were discarded. After discarding the bad trials, an average of 75 % of the unknown and 81 % of the known trials remained for statistical analysis. For each trial, the average pupil diameter during the 200 ms preceding the stimulus onset was subtracted from the task-evoked pupil diameter. Pupil diameters were then converted to millimeters by applying the scale factor. The data were expressed as millimeter deviations from the pretrial baseline. We calculated three dependent variables from the pupillometry data:

Peak dilation: the size of the largest absolute change in pupil size from baseline

Mean change in pupil size: average change in pupil size from baseline across the trial

Maximum percent change in pupillary dilation: the proportion of the peak change in pupil size to baseline pupil size

ERPs The EEG data were preprocessed using EEGLab version 10.2.2 and MATLAB version 8.1. The data were first filtered using a 0.1- to 30-Hz bandpass filter and referenced using an average reference transform to the Cz electrode. Correction for eye movement artifacts was performed by first running a principal component analysis (PCA) on each participant to identify the number of components required to explain 99 % of the data. Independent component analysis (ICA) was then performed using the specified number of components. Following ICA decomposition, eye movements, blinks and other noise components were manually identified and removed from the data.

The resulting cleaned continuous data was segmented into epochs time-locked to the onset of the picture stimulus. Segments extended from 800 ms before to 1,000 ms after the auditory stimulus, in order to include the full response to the picture (presented at -700 ms). Additional bad epochs were identified and rejected using a joint probability computation. The resulting

segments were baseline-corrected using data from the first 100 ms of the segment. In total, an average of 97 % of unknown and 97 % of known trials were included in the statistical analysis.

For the purposes of statistical analysis, the electrodes were broken into six topographic regions across the scalp, including left and right clusters for the frontal, central, and parietal regions (see Fig. 2). The data from these clusters were collapsed over all electrodes. An N400 window of interest was determined on the basis of latency expectations derived from the literature, visual inspection of the waveforms, and running t tests. For running t tests, the raw data were collapsed into 24-ms bins with 12 ms overlap, and the average amplitudes were compared between congruencies within each bin by using paired-sample t tests. More than five bins in a row showing significant differences between conditions ($p < .05$) was deemed a significant window. On the basis of these methods, N400 congruency effects were examined across the window of 450–700 ms post-auditory-onset. We first performed a 2 (knowledge: known, unknown) \times 2 (congruency: congruent, incongruent) \times 3 (site: frontal, central, parietal) \times 2 (hemisphere: left, right) overall ANOVA for the window extending from 450 to 700 ms after sound presentation; significant main effects or interactions were then followed up with additional ANOVAs and post-hoc t tests. For the sake of clarity, we only report and follow up significant ($p < .05$) main effects or interactions.

Results

Behavioral data

Forced-choice task Across all trials, participants selected the correct (named) picture on 79.7 % of trials. Participants had significantly higher accuracy for known trials (99.9 %) than for unknown trials (59.5 %), $t(22) = 16.07$, $p < .0001$, $d = 3.35$. Participants took an average of 1974.8 ms to make their selection across all trials. They were significantly faster to respond on known trials ($M = 1,370.4$ ms) than on unknown trials ($M = 2,579.3$ ms), $t(18) = 17.76$, $p < .0001$, $d = 4.07$.

Congruity task Across all trials, participants made a correct response on auditory/picture congruity on 65.7 % of trials. To compare accuracy on the congruity task, a 2 (congruency: congruent/incongruent) \times 2 (knowledge: known/unknown) repeated measures ANOVA was run on the mean accuracy for each participant. There was a significant interaction of congruency and knowledge [$F(1, 22) = 72.51$, $p < .0001$, $\eta^2 = .74$]. Post-hoc paired-samples t tests showed significantly higher accuracies for unknown incongruent (86.4 %) than unknown congruent trials (33.2 %), $t(22) = 8.54$, $p < .0001$, $d = 1.78$; this pattern may reflect a bias on unknown trials toward responding that the auditory cue did not match the picture. We found no accuracy differences between known

congruent (98.2 %) and known incongruent trials (98.8 %), $t(22) = 1.53$, $p = .14$, $d = 0.32$.

To compare RTs on the congruity task, a 2 (congruency: congruent/incongruent) \times 2 (knowledge: known/unknown) repeated measures ANOVA was run on the mean RTs. There was a significant interaction of knowledge and congruency [$F(1, 22) = 7.21$, $p < .05$, $\eta^2 = .002$]. Post-hoc paired-samples t tests showed significantly slower RTs for unknown congruent trials ($M = 927.8$ ms) than for unknown incongruent trials ($M = 878.7$ ms), $t(22) = 2.69$, $p < .05$, $d = 0.56$. No differences in RTs was apparent between known congruent trials ($M = 759.0$ ms) and known incongruent trials ($M = 793.0$ ms), $t(22) = 1.53$, $p = .14$, $d = 0.32$. That participants responded faster to congruent trials for known words, but to incongruent trials for unknown words, may again reflect a response bias in the incongruent condition.

Word familiarity ratings Known words were given significantly higher word familiarity ratings ($M = 8.99$) than were unknown words ($M = 2.58$), $t(22.05) = 31.35$, $p < .0001$, $d = 6.54$.

EM data

Table 1 presents the mean values on all of the dependent measures derived from eye movements for the known and unknown conditions. Sample eye movement data are shown

Table 1 Means and standard deviations for each dependent variable in the eye movement monitoring and pupillometry data

Dependent Variable	Known		Unknown	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Eye Movements				
Total number of fixations in trial	3.54	0.86	6.98	1.44
Mean fixation duration (ms)	416.4	146.8	354.7	147.5
First fixation duration (ms)	406.8	126.4	328.4	64.7
First dwell on stimulus (ms)	605.4	236.5	445.1	84.5
Latency to first fixation (ms)	742.7	103.2	1,045.0	221.2
Latency to first refixation (ms)	894.7	313.2	1,573.4	594.6
Proportion of fixation duration on stimulus (%)	63.87	17.00	34.14	8.99
Proportion of dwell time on stimulus (%)	42.14	9.61	24.03	4.73
Percentage of trials first fixated (%)	33.50	13.32	27.58	8.38
Percentage of trials last fixated (%)	90.76	8.13	44.98	7.55
Pupillary Dilation				
Peak dilation (mm)	5.43	1.51	7.53	2.30
Mean change in pupil size (mm)	0.01	0.70	1.31	0.74
Max percent change in pupillary dilation (%)	15.78	3.82	22.10	5.45

in Fig. 3. We observed a greater total number of fixations for unknown than for known trials, $t(22) = 17.71$, $p < .0001$, $d = 3.69$. On average, mean fixation durations on the named picture were longer in the known than in the unknown condition, $t(22) = 2.39$, $p < .05$, $d = 0.50$. The length of the first fixation on the named picture was longer for known than for unknown trials, $t(22) = 3.68$, $p < .01$, $d = 0.77$. The length of the first dwell on the named picture was also longer for the known than for the unknown condition, $t(22) = 3.86$, $p < .001$, $d = 0.80$. The latencies to first fixation on the named picture, $t(22) = 8.56$, $p < .0001$, $d = 1.79$, and to refixation, $t(22) = 8.64$, $p < .0001$, $d = 1.80$, were both shorter for known than for unknown trials. The proportion of time spent fixating on the stimulus (i.e., proportion of fixation duration on the stimulus) was greater in the known than in the unknown condition, $t(22) = 12.99$, $p < .0001$, $d = 2.71$. The proportion of time spent dwelling on the stimulus (i.e., looking at the named picture, with or without fixation) was also greater in the known than in the unknown condition, $t(22) = 11.13$, $p < .0001$, $d = 2.32$. The stimulus was the first picture to be fixated on a significantly higher percentage of known than of unknown trials, $t(22) = 2.46$, $p < .05$, $d = 0.51$. Finally, the stimulus was also the last picture to be fixated on a significantly higher percentage of known than of unknown trials, $t(22) = 19.55$, $p < .0001$, $d = 4.08$.

Pupillometry data

The pupillometry data are also summarized in Table 1. Larger peak dilations (relative to baseline) were observed for unknown than for known trials, $t(22) = 9.24$, $p < .0001$, $d = 1.93$. Mean changes in pupil size from baseline were also larger for unknown than for known trials, $t(22) = 8.32$, $p < .0001$, $d = 1.73$. The maximum percent change in pupillary dilation was also larger for unknown than for known trials, $t(22) = 10.86$, $p < .0001$, $d = 2.26$.

ERP data

ERPs for the four conditions, as well as topographical plots of the incongruent–congruent difference for the known and unknown conditions, are presented in Fig. 4.

A 2 (congruency: congruent/incongruent) \times 2 (knowledge status: known/unknown) \times 3 (site: frontal/central/parietal) \times 2 (hemisphere: left/right) repeated measures ANOVA was performed on the average amplitudes over a window from 450 to 700 ms after sound presentation (shaded regions in Fig. 4). The full results can be found in Table 2. We observed a significant three-way interaction of knowledge, congruency, and site [$F(2, 44) = 6.14$, $p < .01$, $\eta^2 = .07$].

To explore this interaction, we performed a 2 (congruency) \times 2 (knowledge) ANOVA for each site (collapsed over hemisphere). There was a significant interaction of congruency and

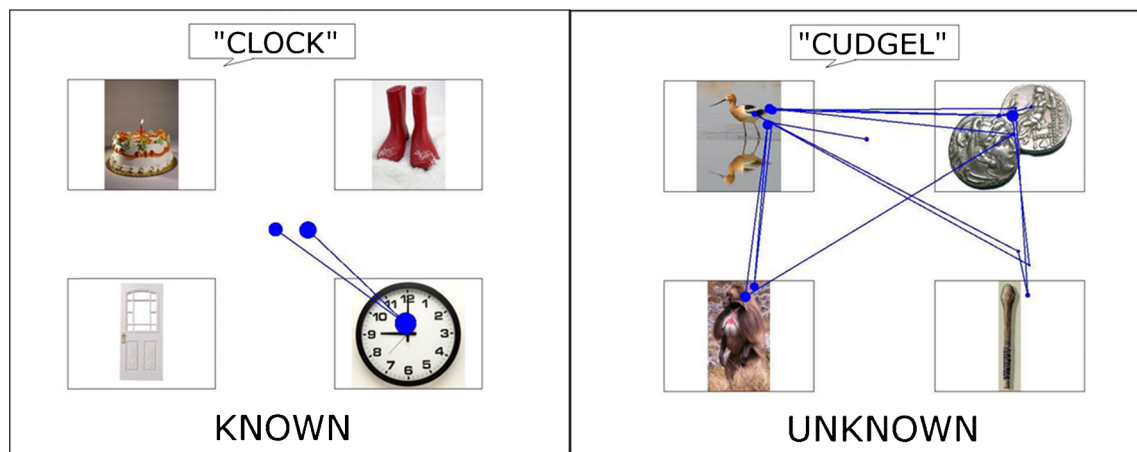


Fig. 3 Sample eye movement patterns for a single participant for a single trial in the known (left) and unknown (right) conditions. Blue dots indicate fixations (with the size of the dot indicating the length of the fixation)

knowledge at parietal sites [$F(1, 22) = 7.34, p < .05, \eta^2 = .02$]. We investigated this interaction at parietal sites by performing a two-way (congruency) ANOVA separately for the known and unknown words. For the known condition, a main effect of congruency emerged [$F(1, 22) = 7.69, p < .05, \eta^2 = .07$]: The mean amplitude observed to known congruent items ($M = 0.08 \mu V, SE = 0.07$) was more positive than that observed to known incongruent items ($M = -0.21 \mu V, SE = 0.10$). No such difference by congruency was observed for unknown items ($F < 1, p = .39, \eta^2 = .002$).

To summarize, a significant N400 congruency effect (a reduction in the amplitude of the N400 for congruent trials, relative to incongruent trials) occurred from 450 to 700 ms over bilateral parietal electrode locations—but only for the known items. No such N400 congruency effect was found for the unknown items.

Effects of familiarity observed to the picture In addition to the expected N400 congruency effect, we also observed an earlier difference in the waveforms recorded to the picture, before the auditory stimulus was presented. Because the auditory word had not yet been presented, any difference observed in this time window would be tied to knowledge differences for the pictures themselves, and could not be linked to congruity (since this was determined by the auditory stimulus). To examine this difference further, we collapsed the ERPs across congruence conditions to look at the differences elicited to the pictures in the known and unknown conditions; these ERPs are shown in Fig. 5.

Running *t* tests identified a sustained difference between the known and unknown conditions beginning approximately 200 ms after picture onset (i.e., -600 ms, relative to the onset of the auditory stimulus). The length of this significant window differed over sites. To compare these effects statistically, a 2 (knowledge: known/unknown) \times 3 (site: frontal/central/parietal) \times 2 (hemisphere: left/right)

repeated measures ANOVA was run on the mean amplitudes for the known and unknown conditions (collapsed over congruency) over a window from 200 to 500 ms after picture presentation (-500 to -200 ms, relative to onset of the auditory token). This window was chosen as the minimum length at which all sites showed differences in the running *t* tests (Fig. 5). The ANOVA showed an interaction of knowledge and site [$F(2, 44) = 28.41, p < .0001, \eta^2 = .07$; see Table 3 for the full results]. To follow up this interaction, we collapsed over hemispheres and performed a two-way (knowledge) ANOVA for each site. Frontal sites showed a main effect of knowledge [$F(1, 22) = 30.00, p < .0001, \eta^2 = .01$], such that the mean amplitude to the known condition ($M = -0.93 \mu V, SE = 0.16$) was more positive than that to the unknown condition ($M = -1.19 \mu V, SE = 0.18$). This effect was also evident over central sites, where there was also a main effect of knowledge [$F(1, 22) = 11.91, p < .01, \eta^2 = .01$] due to a greater relative positivity to the known ($M = -0.24 \mu V, SE = 0.10$) than to the unknown ($M = -0.39 \mu V, SE = 0.12$) condition. At parietal sites, we also found a main effect of knowledge [$F(1, 22) = 31.89, p < .0001, \eta^2 = .02$], but here, the polarity of the effect was reversed: There was a greater relative positivity to unknown ($M = 1.20 \mu V, SE = 0.14$) than to known ($M = 0.97 \mu V, SE = 0.13$) items.

Thus, we observed differences in the response to the visual stimulus between the known and unknown conditions prior to the presentation of the (congruent or incongruent) auditory stimulus. At frontal and central electrode locations, this difference was in the form of a relatively more positive mean amplitude to pictures in the known condition, whereas at parietal sites, there was a greater relative positivity to pictures in the unknown condition. This difference onset early across all scalp locations (beginning approximately 200 ms after presentation of the picture) and extended in time for several hundred milliseconds (especially at parietal sites).

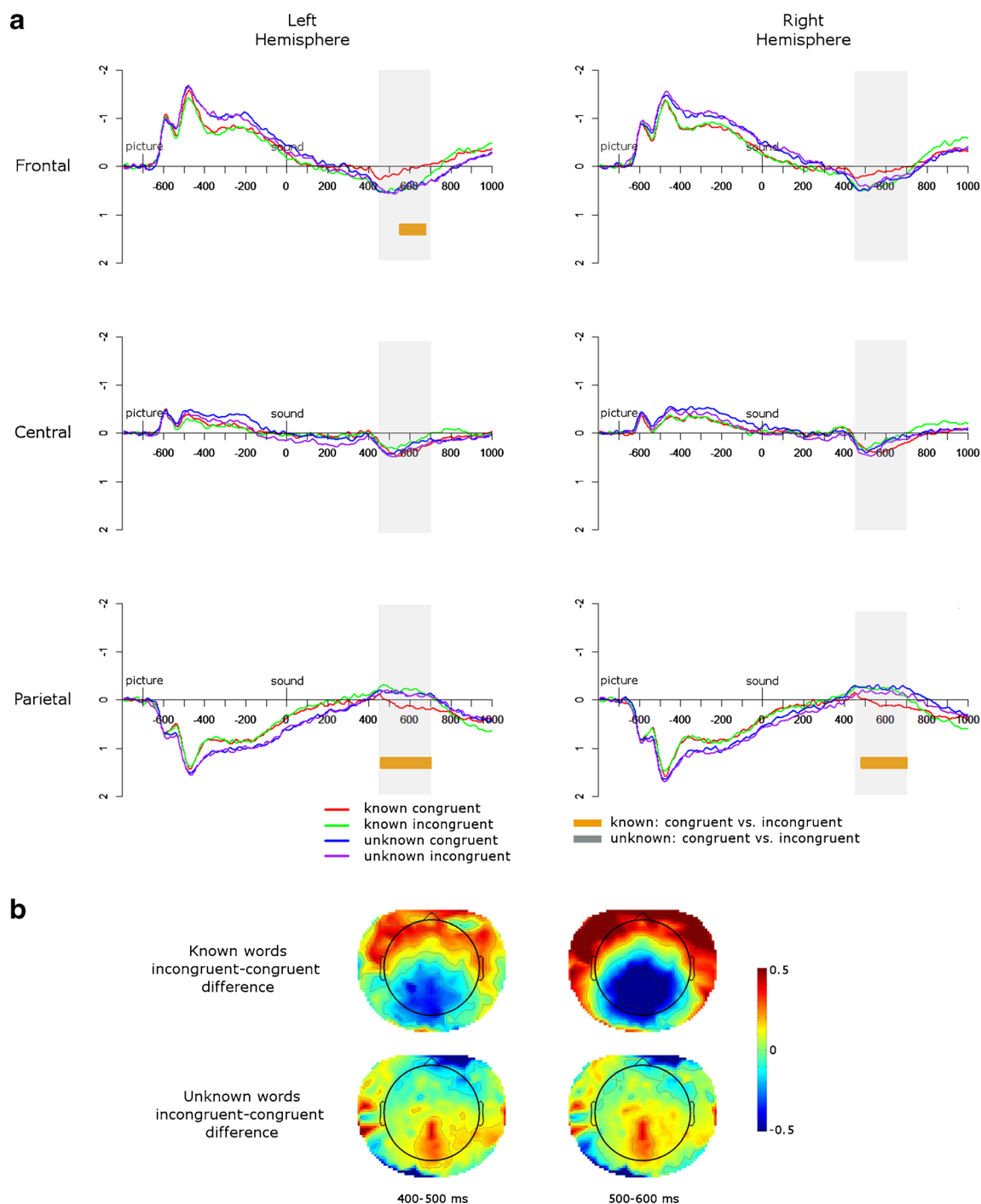


Fig. 4 (a) N400 effects of picture-word congruency for known and unknown words. ERPs are grand averages across all participants, collapsed across electrodes in the frontal, central, and parietal regions. Gray shading indicates the N400 window (450–700 ms after the onset of

the auditory stimulus). (b) Topographic distribution of the difference waves (incongruent – congruent) for the known and unknown conditions across two time windows

Correlations among measures

To consider the relationship between knowledge status and our various dependent measures, we ran Pearson's correlations between several variable pairs separately for known and unknown items.

Behavioral measures with implicit measures First, we examined the correlations between the three behavioral measures and the implicit measures; these results are shown in Table 4. The first behavioral measure was the PPVT score for each participant. For known items, we observed several negative correlations between PPVT score and the EM duration

Table 2 Results from the 2 (knowledge) \times 3 (congruency) \times 3 (site) \times 2 (hemisphere) analysis of variance on the average amplitudes over a window from 450 to 700 ms after sound presentation

Main effect or interaction	<i>F</i> Value	<i>df</i>	<i>p</i> Value	η^2 Value
Knowledge	<1	1, 22	.57	.002
Congruency	<1	1, 22	.85	.0002
Site	3.37	2, 44	<.05*	.11
Hemisphere	<1	1, 22	.43	.01
Knowledge \times Congruency	1.30	1, 22	.27	.008
Knowledge \times Site	2.31	2, 44	.10 [§]	.03
Congruency \times Site	3.89	2, 44	<.05*	.04
Knowledge \times Hemisphere	2.50	1, 22	.13	.004
Congruency \times Hemisphere	<1	1, 22	.98	<.0001
Site \times Hemisphere	<1	2, 44	.96	.0002
Knowledge \times Congruency \times Site	6.14	2, 44	<.01**	.07
Knowledge \times Congruency \times Hemisphere	<1	1, 22	.78	.0002
Knowledge \times Site \times Hemisphere	1.52	2, 44	.23	.001
Congruency \times Site \times Hemisphere	3.95	2, 44	<.05*	.005
Knowledge \times Congruency \times Site \times Hemisphere	<1	2, 44	.79	.0003

Significant effects are indicated: [§] trend, $p < .10$, * $p < .05$, ** $p < .01$

measures (such as mean fixation duration, first fixation duration, and proportion of dwell time on the stimulus), suggesting that larger vocabulary scores were associated with shorter looking times for known items. For unknown words, on the other hand, such correlations were not observed; the only significant correlation for this set was a positive correlation between PPVT score and the percentage of trials on which the named item was the last to be fixated.

The second behavioral measure was the RT on the forced-choice task. We ran correlations between these RTs and the EM and pupillometry measures, which were collected using the same paradigm. These are also shown in Table 4. Of note, for known words, we observed several positive correlations between the RT and EM measures, suggesting that longer times to select the named picture from the display were accompanied by longer looking times. This was not seen for unknown items, for which we saw only one marginally significant correlation between RT and the EM measure of first dwell time. There was, however, a significant negative correlation between RT and the mean change in pupil size for unknown items, suggesting that faster RTs were accompanied by smaller changes in pupil size for unknown items.

The third behavioral measure was RT on the congruity task, which we correlated with the N400 effect size from the concurrent ERP task. For the congruity task, RT effect sizes were calculated by subtracting the mean RT in the congruent condition from the mean RT in the incongruent condition, separately for known and unknown items. From the ERP data, the

N400 effect was calculated by first calculating the difference wave (incongruent minus congruent) for each individual word, then finding the most negative peak in the difference wave within a window from 200 to 800 ms after sound presentation. The average difference wave amplitude within a 200-ms window around the most negative difference wave peak was then calculated, yielding an *N400 effect* measure for each individual word, which was averaged over known and unknown trials. As can be seen in Table 4, no significant correlations emerged between the RT effect size on the congruity task and N400 effect size.

Correlations among implicit measures We also ran correlations among the various implicit measures themselves. The results of these correlations are shown in Tables 5 (for known items) and 6 (for unknown items). Some patterns are worth highlighting. First, for both known and unknown words, the EM measures are all highly intercorrelated, as are two of the three pupillometry measures (peak dilation and maximum percent change in pupil dilation), suggesting that these measures may be tapping into the same underlying processes. There are also a number of significant correlations between the measures from the different implicit assessment techniques; for example, for known words (but not for unknown words), we observed significant correlations between the N400 effect size and EM measures such as the mean fixation duration and first fixation duration.

Discussion

In the present study, we used measures from three different implicit assessment techniques (eye movement monitoring, pupillometry, and event-related potentials) to study receptive vocabulary knowledge in normal adults. Specifically, we looked for differences between these measures to high-frequency, highly familiar words, which were expected to be known by all adult participants, and to low-frequency, unfamiliar words, which were expected to be unknown. The behavioral measures that we administered supported the distinction between these two sets of words. First, offline word familiarity ratings suggested that participants were very familiar with the high-frequency words in our known condition, and were relatively unfamiliar with the low-frequency words in the unknown condition. Additionally, on both the forced-choice and congruity tasks, participants were more accurate and faster when making responses to known than to unknown items. These results support the distinction between the two sets of words and suggest that these are appropriate sets in which to look for processing differences using our implicit measures.

The ERP technique has previously proven useful in detecting differences in receptive vocabulary knowledge for known

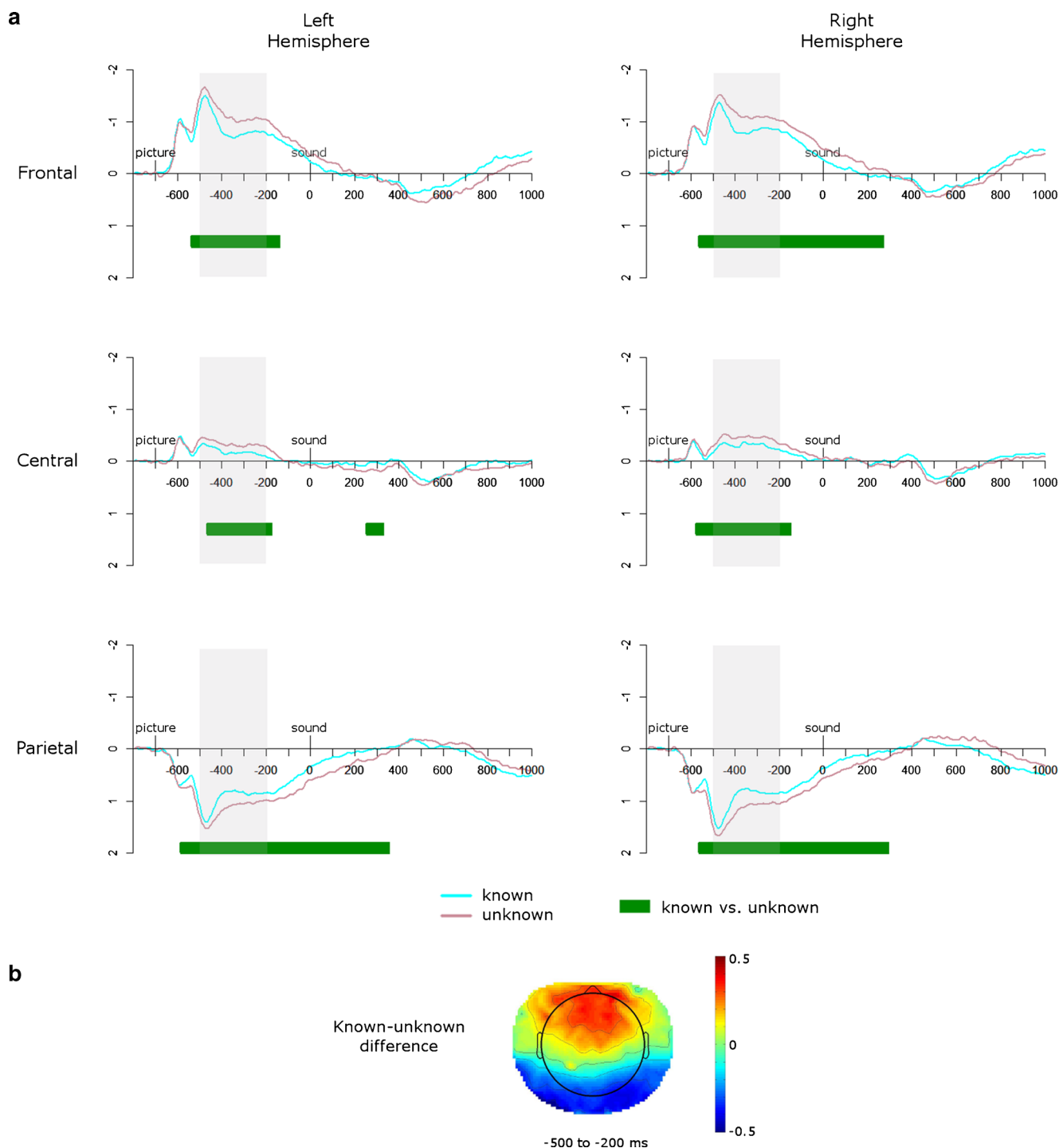


Fig. 5 (a) Effects of picture familiarity for known and unknown words. ERPs are averaged over congruent and incongruent conditions for known and unknown stimuli, and collapsed across electrodes in the frontal,

central, and parietal regions. (b) Topographic distribution of the difference waves (known – unknown) across the time window of interest (–500 to –200 ms before sound presentation: gray shading)

and unknown words in a variety of participant groups. Our results with ERPs replicated those of several other research groups using a similar congruency paradigm (Byrne et al., 1999; Connolly & D'Arcy, 2000; Connolly et al., 2000; D'Arcy et al., 2003; Friedrich & Friederici, 2004, 2005a, b,

2010; Henderson et al., 2011; Marchand et al., 2002; Torkildsen et al., 2008). A reliable reduction in the amplitude of the N400 was observed in congruent versus incongruent word/picture pairings, but only for the items that were expected to be known to the participants. For the unknown word

Table 3 Results from the 2 (knowledge) \times 3 (site) \times 2 (hemisphere) analysis of variance on the known and unknown average amplitudes (collapsed over congruency) over a window from -500 to -200 ms before sound presentation

Main Effect or Interaction	<i>F</i> Value	<i>df</i>	<i>p</i> Value	η^2 Value
Knowledge	9.53	1, 22	<.01**	.006
Site	25.7	2, 44	<.0001***	.49
Hemisphere	<1	1, 22	.79	.0007
Knowledge \times Site	28.41	2, 44	<.0001***	.07
Knowledge \times Hemisphere	<1	1, 22	.72	.0001
Site \times Hemisphere	1.49	2, 44	.24	.006
Knowledge \times Site \times Hemisphere	<1	2, 44	.72	<.0001

Significant effects are indicated: ** $p < .01$, *** $p < .001$

pairings, no difference in the amplitude of the N400 was observed. The observed congruency effect for the known words would be expected if participants are able to use their knowledge of word meanings to use the picture as context when deciding whether the accompanying auditory token is congruent or incongruent. However, because such underlying semantic knowledge is not available in the case of the unknown words, the ease of processing that results from a congruency between a word and its context (and the resultant reduction of the N400) cannot occur. The N400 effect observed in our experiment was slightly delayed in latency relative to canon-

ical N400 effects; this is likely due to the fact that we time-locked our ERPs to the onset of the auditory word. Processing may not have fully engaged until the offset of this word, or at least until enough of the word had been presented to allow for lexical selection.

In addition to the replication of the congruency effect for known (but not for unknown) words, we observed differences in the ERPs elicited by the visual stimulus itself, prior to the presentation of the auditory stimulus that determined congruency. Differences (in the form of a greater relative positivity to known pictures at frontal and central sites, and a greater relative positivity to unknown pictures at parietal electrode sites) were observed shortly after the onset of the picture stimulus, and remained over a period of several hundred milliseconds, especially at parietal sites. Because the auditory stimulus had not yet been presented, these differences were instead elicited by the pictures themselves, and may thus reflect differences in the familiarity with known items relative to unknown items. For example, this pattern could be interpreted as an N400-like effect of semantic integration, such that increased difficulty with integrating the unknown pictures into semantic memory elicited the observed patterns. However, if this were the case, this would predict an increased negativity to unknown pictures over parietal sites, whereas we currently see an increased positivity in this region. Although this effect and its interpretation need further replication, the finding of familiarity differences to the pictures themselves would be a potentially

Table 4 Correlations between implicit measure dependent variables and behavioral measures

	PPVT		Forced-Choice RT		Congruity RT	
	Known	Unknown	Known	Unknown	Known	Unknown
EM						
Total number of fixations	-.13	-.27	.41 [§]	.29		
Mean fixation duration	-.49*	-.29	.57*	.09		
First fixation duration	-.44*	-.13	.56*	.12		
First dwell time on stimulus	-.34	.09	.91**	.43 [§]		
Latency to first fixation	.15	-.01	.42 [§]	.11		
Latency to first refixation	-.26	-.36 [§]	-.17	.07		
Proportion of fixation duration on the stimulus	-.35 [§]	.02	.49*	.14		
Proportion of dwell time on stimulus	-.42*	.17	.69**	.06		
First (%)	.02	-.04	-.23	-.06		
Last (%)	-.14	.54**	.37	-.30		
PD						
Peak dilation	.02	.11	-.29	-.07		
Mean change in pupil size	-.14	.28	-.26	-.72**		
Maximum percent change in pupil size	-.04	.10	-.14	-.01		
ERP						
N400 effect	-.37 [§]	-.33			.29	.11

Behavioral RTs are correlated with dependent measures that were collected during the same paradigm only. The congruity RT and N400 effects were based on incongruent – congruent differences; see the text for further explanation. Statistically significant correlations are indicated with asterisks: [§] $p < .10$, * $p < .05$, ** $p < .01$

Table 5 Correlations between all eye movement monitoring (EM), pupillometry, and event-related potential (ERP) measures, for known trials only

Known trials	Total number of fixations	Mean fixation duration	First fixation duration	First dwell time	Latency to first fixation	Latency to first refixation	Proportion of fixation duration on the stimulus	Proportion of dwell time on stimulus	First (%)	Last (%)	Peak dilation	Mean change in pupil size	Maximum percent change in pupil size	N400 effect
EM measures	1													
Mean fixation duration	.11	1												
First fixation duration	-.03	.96**	1											
First dwell time	.37 [§]	.74**	.66**	1										
Latency to first fixation	-.46*	.13	.27	.22	1									
Latency to first refixation	-.15	.18	.27	-.07	.18	1								
Proportion of fixation duration on the stimulus	.12	.59**	.49*	.60**	.04	-.05	1							
Proportion of dwell time on stimulus	.61**	.67**	.49	.85**	-.29	-.20	.62**	1						
First (%)	-.69**	-.12	-.01	-.20	.53**	.30	.06	-.52*	1					
Last (%)	.44*	.34	.08	.45*	-.34	-.37 [§]	.46*	.71**	-.40 [§]	1				
Peak dilation	-.51*	-.04	.06	-.35	.21	.07	-.09	-.41	.28	-.07	1			
Mean change in pupil size	-.18	-.13	-.11	-.21	.02	.35	-.14	-.23	.49*	-.18	.17	1		
Maximum percent change in pupil size	-.65**	.15	.26	-.12	.33	.18	.19	-.28	.56**	-.10	.81**	.25	1	
ERP measure	.34	.43*	.45*	.39 [§]	.21	.31	.08	.30	-.20	-.01	-.15	-.11	-.22	1

Statistically significant correlations are indicated with asterisks: [§] $p < .10$, * $p < .05$, ** $p < .01$.

Table 6 Correlations between all eye movement monitoring (EM), pupillometry, and event-related potential (ERP) measures, for unknown trials only

Unknown trials		Total number of fixations	Mean fixation duration	First fixation duration	First dwell time	Latency to first fixation	Latency to first refixation	Proportion of fixation on the stimulus	Proportion of dwell time on stimulus	First (%)	Last (%)	Peak dilation	Mean change in pupil size	Maximum percent change in pupil size	N400 effect
EM measures	Total number of fixations	1													
	Mean fixation duration	.09	1												
	First fixation duration	-.03	.88**	1											
	First dwell time	.35§	.29	.39§	1										
	Latency to first fixation	-.34	-.10	.18	-.03	1									
	Latency to first refixation	-.03	.54**	.55**	-.09	.33	1								
	Proportion of fixation on the stimulus	.25	.15	.14	.47*	-.15	-.15	1							
	Proportion of dwell time on stimulus	.51*	.24	.09	.67**	-.64**	-.23	.47*	1						
	First (%)	-.14	-.09	-.13	.05	.03	.07	-.01	.08	1					
	Last (%)	-.13	.06	-.05	.01	-.34	-.01	.01	.46*	-.04	1				
Pupillometry measures	Peak dilation	-.64**	.19	.28	-.21	.25	.13	-.13	-.30	.10	.26	1			
	Mean change in pupil size	-.30	-.15	-.11	.00	-.01	-.20	-.02	.08	.20	-.07	-.05	1		
	Maximum percent change in pupil size	-.70**	.11	.24	-.08	.41§	.22	.05	-.31	.25	.10	.79**	.13	1	
ERP measures		.27	.21	.25	.34	.08	-.08	.17	.08	-.17	-.35	-.25	-.13	-.35	1

Statistically significant correlations are indicated with asterisks: § $p < .10$, * $p < .05$, ** $p < .01$

important addition to the body of research on the use of ERPs to study receptive vocabulary knowledge.

Importantly, the results from our two additional implicit assessment techniques showed similar patterns of reliable differentiation between the processing of known and unknown words. During a forced-choice task, we simultaneously collected EM and pupillometry data. In the EM data, participants' eye movements were generally faster to and more consistently focused on the named picture for known than for unknown words. The fixation duration measures (including mean fixation duration, first fixation duration, and first dwell duration) were all longer for known than for unknown words, demonstrating that participants spent more time looking at the named picture when it was associated with a known word than when it was associated with an unknown word. Participants were faster to move their eyes to the named picture for the first time (latency to first fixation) and to move their eyes back to the named picture after having left it (latency to first refixation) for known than for unknown words. Finally, proportional measures showed a similar finding: Proportions of fixations on the stimulus and proportions of dwell time on the stimulus showed that participants looked more at the named picture, and for longer amounts of time, in the known than in the unknown condition. Thus, reliable differences were observed across all dependent measures of eye movements, suggesting that the processing of known words was different than that of unknown words, in ways that suggest that identifying the named picture was easier for participants in the known condition. Similar findings were observed in the pupillometry data, in which all three of our dependent measures (peak dilation, average change in pupil size, and percent change in pupillary dilation) showed larger changes in the unknown than in the known condition. Because pupillary dilation has been shown to increase with task difficulty across a range of tasks in previous research, these results again support the conclusion that processing in the forced-choice task paradigm was more difficult for words that were expected to be unknown by our participants.

Our ERP results, then, replicated previous findings of reliable differences in processing measures of words depending on their receptive vocabulary status (known or unknown). Our EM and pupillometry results demonstrated that such differences can also be identified using other implicit assessment techniques that, to our knowledge, have not previously been used explicitly to study receptive vocabulary knowledge in any participant group. Thus, the EM and pupillometry results provide important confirmation of the ERP differences that were observed, with the similar advantage of not having to rely on explicit behavioral responses from participants. Crucially, although we collected behavioral measures in the present study, all three of these implicit measures yield reliable differences between known and unknown words that do not rely on these behavioral responses. These implicit techniques

are thus invaluable for studying receptive vocabulary knowledge in populations who do not, or cannot, make overt responses.

We also looked at the correlations between our implicit measures and behavioral measures, as well as the correlations between the various implicit measures themselves. We observed significant relationships between a standardized measure of vocabulary (the PPVT) and several of our EM measures; a marginally significant correlation was also observed between PPVT score and N400 effect size. We also observed significant correlations between RTs on the forced-choice task and a number of EM measures. We did not observe significant relationships between the N400 effect size and the RT difference measure on the congruity task; this may have been due in part to a lack of statistical power, due to the inherently noisy ERP data.

We observed a number of significant correlations between measures from the different assessment techniques, as can be seen, for example, in the positive correlations between N400 effect size and some of the EM duration measures for known words. However, in a number of cases significant correlations were not observed across the measures. In some cases, this may have been due to a lack of statistical power, due to the inherent noisiness of the ERP data; for example, negative correlations between N400 effect size and the pupillary measures (such that larger N400 differences were observed when pupillary dilation changes were smaller, and vice versa) were observed for both the known and unknown words, but these correlations did not reach significance. Other cases, however, may be due to the fact that the implicit measures reflect underlying processes that, though all engaged in the service of word and picture recognition, potentially vary quite widely in their exact cognitive function. These differences may be exacerbated by the differences in task requirements for the forced-choice and congruity paradigms. What is important for the present purposes is that even when the implicit measures did not correlate with each other (perhaps because they were tapping into complementary but nonoverlapping cognitive processes), all three were still capable of differentiating groups of known words from groups of unknown words.

These results suggest that eye movements and pupillometry might provide techniques alternative to ERPs for the assessment of receptive vocabulary knowledge. The availability of alternatives to ERPs for such testing might be appealing for several reasons, as we proposed in the introduction. First, EM and pupillometry might be available to researchers or clinicians for whom the ERP methodology is not (for practical, financial, or other reasons). Being able to use EM or pupillometry might thus make the implicit assessment of receptive vocabulary more accessible to a wider group of individuals for whom such knowledge might prove useful (e.g., for research purposes or for the purposes of developing rehabilitative therapy). Second, EM and pupillometry

recording might be accomplished better with some participant groups than are ERPs. Many eyetracking systems are capable of collecting EM and pupillary data in a noninvasive manner, without the need for equipment that touches the participant in any way. This is not the case for ERP research, which requires the placement of electrodes on participants' scalps. Even under the most ideal circumstances—for example, with modern recording systems that try to minimize the time required for electrode application—the process of applying the electrodes to the scalp, and the necessity of keeping them there during data acquisition, can prove difficult with some participant groups (such as small children and infants), and perhaps can be highly stressful to others (such as individuals with autism, who have demonstrated sensitivities to and varying tolerances for objects placed on their person). Similarly, the desire to minimize EEG artifacts during data acquisition is more easily accomplished with some participant groups than with others; although eyetracking systems have their own constraints in this regard, some participant groups may be better able to comply with the eyetracking restrictions than with those imposed by ERPs.

These results suggest that in the future, the use of eyetracking measures (EM and pupil dilation) might help to overcome one of the greatest obstacles to using ERPs to better understand receptive vocabulary: the need to average the dependent measure across multiple items in ERP experiments. This makes it difficult to examine the event-related signals to individual items, such as vocabulary words. Ideally, though, this might be *exactly* what we would most like to do: to make a determination, on the basis of the response to an individual item, whether that specific item is likely to be known or unknown to the individual participant. Such information would be invaluable to clinicians and teachers in developing and personalizing instruction. The use of EM and pupillometry might offer just such a possibility, since these measures do not depend on signal averaging across trials of like types. In fact, the multiple dependent measures that can be recorded using either EM or pupillometry might have the additional advantage of allowing the trained evaluator to look at patterns across dependent measures for a single vocabulary word to make a determination about whether that word is or is not known to an individual participant. We have explored these possibilities by using ERP, EM, and pupillometry measures to model subjective knowledge ratings using mixed-effects logistic regression (Coderre, Gordon, & Ledoux, under review).

Finally, the use of any one of these implicit measurement techniques (EM, pupillometry, or ERPs) may prove especially advantageous to the study of receptive vocabulary knowledge in participant groups from whom reliable overt behavioral responses are difficult to collect. These very frequently may be the very participant groups in which such knowledge could be most beneficial in terms of further learning. In ongoing work, we have been using these same paradigms to collect

implicit measures of receptive vocabulary knowledge in typically developing children and in high- and low-functioning individuals with autism. This last group, in particular, has been especially difficult to study using more traditional measurement techniques. Given the pervasive language and communication deficits observed for low-functioning individuals with autism, having measures of what words may or may not be understood by individual participants could prove useful in terms of further instruction and in terms of improving caregiver communication.

Author note This research was supported by grants from the Nancy Lurie Marks Foundation, the Department of Defense Autism Research Program (Grant No. AR093137), the Therapeutic Cognitive Neuroscience Fund, and the Benjamin and Adith Miller Family Endowment on Aging, Alzheimer's, and Autism Research. E.B. is now at the University of Rochester; I.G. is now at the University of Wisconsin.

References

- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439. doi:10.1006/jmla.1997.2558
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91, 276–292. doi:10.1037/0033-2909.91.2.276
- Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). New York, NY: Cambridge University Press.
- Brown, C. M., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5, 34–44.
- Bryshaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41, 977–990. doi:10.3758/BRM.41.4.977
- Byrne, J. M., Connolly, J. F., MacLean, S. E., Dooley, J. M., Gordon, K. E., & Beattie, T. L. (1999). Brain activity and language assessment using event-related potentials: Development of a clinical protocol. *Developmental Medicine and Child Neurology*, 41, 740–747.
- Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory*, 92, 225–234.
- Connolly, J. F., & D'Arcy, R. C. N. (2000). Innovations in neuropsychological assessment using event-related brain potentials. *International Journal of Psychophysiology*, 37, 31–47.
- Connolly, J. F., D'Arcy, R. C. N., Newman, R. L., & Kemps, R. (2000). The application of cognitive event-related brain potentials (ERPs) in language-impaired individuals: Review and case studies. *International Journal of Psychophysiology*, 38, 55–70.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology*, 6, 84–107.
- D'Arcy, R. C. N., Marchand, Y., Eskes, G. A., Harrison, E. R., Phillips, S. J., Major, A., & Connolly, J. F. (2003). Electrophysiological assessment of language function following stroke. *Clinical Neurophysiology*, 114, 662–672.
- Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access:

- Evidence for lexical competition. *Language and Cognitive Processes*, 16, 507–534.
- Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken word recognition. *Psychonomic Bulletin & Review*, 12, 453–459. doi:10.3758/BF03193787
- Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica*, 30, 412–431. doi:10.1016/0001-6918(69)90065-1
- Dunn, L. M., & Dunn, L. M. (2007). *Peabody picture vocabulary tests* (4th ed.). Circle Pines, MN: American Guidance Service.
- Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, 24, 409–436.
- Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, 146, 668–670.
- Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, 42, 98–116.
- Fernald, A., Swingle, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, 72, 1003–1015.
- Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, 16, 1465–1477. doi:10.1162/0898929042304705
- Friedrich, M., & Friederici, A. D. (2005a). Lexical priming and semantic integration reflected in the ERP of 14-month-olds. *NeuroReport*, 16, 653–656.
- Friedrich, M., & Friederici, A. D. (2005b). Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, 17, 1785–1802. doi:10.1162/089892905774589172
- Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *NeuroReport*, 19, 1757–1762.
- Friedrich, M., & Friederici, A. D. (2010). Maturing brain mechanisms and developing behavioral language skills. *Brain and Language*, 114, 66–71.
- Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin*, 77, 340–355. doi:10.1037/h0032456
- Henderson, L. M., Baseler, H. A., Clarke, P. J., Watson, S., & Snowling, M. J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary, and decoding. *Brain and Language*, 117, 88–99.
- Hess, E. H. (1965). Attitude and pupil size. *Scientific American*, 212, 46–54.
- Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, 143, 1190–1192.
- Hess, E. H., Seltzer, A. L., & Shlien, J. M. (1965). Pupil response of hetero- and homosexual males to pictures of men and women: A pilot study. *Journal of Abnormal Psychology*, 70, 165–168.
- Hoeks, B., & Levelt, W. J. M. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. *Behavior Research Methods, Instruments, & Computers*, 25, 16–26. doi:10.3758/BF03204445
- Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, 30, 47–61.
- Huetting, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, 96, B23–B32. doi:10.1016/j.cognition.2004.10.003
- Huetting, F., & Altmann, G. T. M. (2011). Looking at anything that is green when hearing “frog”: How object surface colour and stored object colour knowledge influence language-mediated overt attention. *Quarterly Journal of Experimental Psychology*, 64, 122–145.
- Junge, C., Cutler, A., & Hagoort, P. (2012). Electrophysiological evidence of early word learning. *Neuropsychologia*, 50, 3702–3712.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441–480. doi:10.1016/0010-0285(76)90015-3
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154, 1583–1585.
- Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during a mental task. *Science*, 157, 218–219.
- Karatekin, C., Marcus, D. J., & Couperus, J. W. (2007). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. *Psychophysiology*, 44, 128–144.
- Key, A. P. F., Molfese, D. L., & Ratajczak, E. D. (2006). ERP indicators of learning in adults. *Developmental Neuropsychology*, 29, 379–395.
- Kuipers, J.-R., & Thierry, G. (2011). N400 amplitude reduction correlates with an increase in pupil size. *Frontiers in Human Neuroscience*, 5, 61. doi:10.3389/fnhum.2011.00061
- Kuipers, J.-R., & Thierry, G. (2013). ERP-pupil size correlations reveal how bilingualism enhances cognitive flexibility. *Cortex*, 49, 2853–2860. doi:10.1016/j.cortex.2013.01.012
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4, 463–470.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205. doi:10.1126/science.7350657
- Loewenfeld, I. E. (1993). *The pupil: Anatomy, physiology, and clinical applications* (Vol. 1). Ames, IA: Iowa State University Press.
- Marchand, Y., D’Arcy, R. C. N., & Connolly, J. F. (2002). Linking neurophysiological and neuropsychological measures for aphasia assessment. *Clinical Neurophysiology*, 113, 1715–1722.
- Myung, J.-Y., Blumstein, S. E., Yee, E., Sedivy, J. C., Thompson-Schill, S. L., & Buxbaum, L. J. (2010). Impaired access to manipulation features in apraxia: Evidence from eyetracking and semantic judgment tasks. *Brain and Language*, 112, 101–112. doi:10.1016/j.bandl.2009.12.003
- Myung, J.-Y., Blumstein, S. E., & Sedivy, J. C. (2006). Playing on the typewriter, typing on the piano: Manipulation knowledge of objects. *Cognition*, 98, 223–243. doi:10.1016/j.cognition.2004.11.010
- Odekar, A., Hallowell, B., Kruse, H., Moates, D., & Lee, C.-Y. (2009). Validity of eye movement methods and indices for capturing semantic (associative) priming effects. *Journal of Speech, Language, and Hearing Research*, 52, 31–48.
- Ojima, S., Nakamura, N., Matsuba-Kurita, H., Hoshino, T., & Hagiwara, H. (2010). Neural correlates of foreign-language learning in childhood: A 3-year longitudinal ERP study. *Journal of Cognitive Neuroscience*, 23, 183–199.
- Posner, M. I. (2005). Timing the brain: Mental chronometry as a tool in neuroscience. *PLoS Biology*, 3(e51), 204–206. doi:10.1371/journal.pbio.0030051
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. doi:10.1037/0033-2909.124.3.372
- Swingle, D., & Fernald, A. (2002). Recognition of words referring to present and absent objects in 24-month-olds. *Journal of Memory and Language*, 46, 39–56.
- Tanenhaus, M. K. (2007). Eye movements and spoken language processing. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 443–469). Amsterdam, The Netherlands: Elsevier.
- Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, 29, 557–580.

- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634. doi:[10.1126/science.7777863](https://doi.org/10.1126/science.7777863)
- Torkildsen, J. V. K., Friis Hansen, H., Svangstu, J. M., Smith, L., Simonsen, H. G., Moen, I., & Lindgren, M. (2009). Brain dynamics of word familiarization in 20-month-olds: Effects of productive vocabulary size. *Brain and Language*, 108(1), 73–88. doi:[10.1016/j.bandl.2008.09.005](https://doi.org/10.1016/j.bandl.2008.09.005)
- Torkildsen, J. V. K., Svangstu, J. M., Friis Hansen, H., Smith, L., Simonsen, H. G., Moen, I., & Lindgren, M. (2008). Productive vocabulary size predicts event-related potential correlates of fast mapping in 20-month-olds. *Journal of Cognitive Neuroscience*, 20, 1266–1282. doi:[10.1162/jocn.2008.20087](https://doi.org/10.1162/jocn.2008.20087)
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73, 89–134.
- Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition*, 19, 95–112.
- Yee, E., Huffstetler, S., & Thompson-Schill, S. L. (2011). Function follows form: Activation of shape and function features during object identification. *Journal of Experimental Psychology: General*, 140, 348–363. doi:[10.1037/a0022840](https://doi.org/10.1037/a0022840)
- Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1–14. doi:[10.1037/0278-7393.32.1.1](https://doi.org/10.1037/0278-7393.32.1.1)