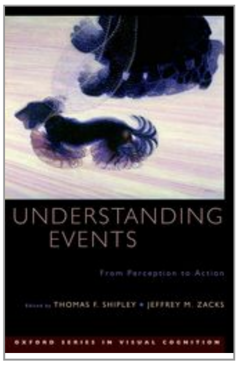

Understanding Events: From Perception to Action
Thomas F. Shipley and Jeffrey M. Zacks

Print publication date: 2008
Print ISBN-13: 9780195188370
Published to Oxford Scholarship Online: May 2008
DOI: 10.1093/acprof:oso/9780195188370.001.0001

Neurocognitive Mechanisms of Human Comprehension

Tatiana Sitnikova
Phillip J. Holcomb
Gina R. Kuperberg

DOI:10.1093/acprof:oso/9780195188370.003.0026

Abstract and Keywords

This chapter begins with a discussion of evidence for distinctions between two semantic comprehension systems in the language domain: a system that maps the perceived information on graded semantic representations and a system that utilizes particular semantic requirements of verbs. It then reviews similar research using static and motion pictures. It argues that the two mechanisms of language comprehension might be analogous to the systems that use graded semantic representations and action-based requirements to make sense of the visual world. The experiments that are reviewed in this chapter examine questions of both how comprehenders understand relationships between the elements within individual events and how they understand the relationships between events. Experiments that have used event-related potentials (ERPs) are also highlighted.

Keywords: language comprehension, visual world comprehension, event-related potentials, semantic comprehension systems, static pictures, motion pictures

As humans make sense of the world, such as when processing language or watching events unfold around them, they must combine separate aspects of the incoming stimuli into a coherent gestalt of overall meaning. Comprehending an individual event depends on determining the nature of its central action and the roles (often termed "thematic roles") played by the people and objects in this action (Klix, 2001; Nowak, Plotkin, & Jansen, 2000). For example, at a birthday party, an observer of an event “woman cutting a cake with a knife” would understand that the action is “cutting,” that the woman plays the role of an Agent (the person who is doing the cutting), that the cake is a Patient (the object that is being cut), and that the knife is an Instrument (the object that is used for cutting). Moreover, each event must be understood in the context of the preceding events (Klix, 2001; Knutson, Wood, & Grafman, 2004; Wood & Grafman, 2003). For example, if, before the cake is cut, one of the children at the party blows out the candles on the cake, the observer would understand the relationships between these two events: at a birthday party, blowing out the candles on a cake is a prerequisite for cutting it. This chapter reviews research suggesting that there are two separate **(p.640)** neurocognitive mechanisms mediating such comprehension processes. The first mechanism relies on relationships of various strengths stored in comprehenders’ semantic memory of the world (we term this knowledge “graded semantic representations”). The second mechanism relies on discrete, rule-like knowledge of what is necessary for the real-world actions (we term this knowledge “action-based requirements”).

The notion of structured semantic memory representations has a long history in the theory of comprehension. Individual concepts in semantic memory are thought to have connections of various strengths, depending on factors such as their feature similarity or how often they have been experienced in the same context (Fischler & Bloom, 1985; Hutchison, 2003; Meyer & Schvaneveldt, 1971; Neely, 1991; Rosch, 1975; Rosch & Mervis, 1975; Stanovich & West, 1983). These graded semantic representations are continuously accessed and used as comprehension takes place in real time (online), being especially useful in familiar circumstances. Perceiving only a few details allows comprehenders to access representations of the related concepts and, as a result, to rapidly grasp the likely overall meaning of everyday situations and to prepare for what would be expected to come next. By mapping the perceptual input on these graded semantic representations, comprehenders can build expectations at various levels ranging from specific entities that are likely to play a given role in a real-world action (Ferretti, McRae, & Hatherell, 2001; McRae, Hare, Elman, & Ferretti, 2005) to the probable spatiotemporal relationships between individual events (Abelson, 1981; Bower, Black, & Turner, 1979; Schank & Abelson, 1977; van der Meer, Beyer, Heinze, & Badel, 2002). For example, the representational network of a conventional birthday party scenario would include strong associations between the Agent role of the cutting action and such features as <adult> and <able to perform volitional actions>, between the Patient role and such features as <has frosting> and <unsturdy>, and between the Instrument role and such features as <has handle> and <has a sturdy, sharp edge>. At a more global level, this information would be linked to the events that usually precede and follow cutting the cake at a birthday party. During comprehension, as one views, for instance, a boy blowing out candles on a birthday cake, accessing the related representations in semantic memory would allow this observer to anticipate that the boy's parent would soon be using a knife to cut the cake.

(p.641) Such graded semantic representations, however, are relatively rigid in that what is stored is descriptive in nature, without regard as to whether any given component is *necessary* for a particular action. As a result, accessing these representations has limited utility for comprehending unfamiliar and unusual situations, and can not readily account for humans' remarkable ability to make sense of such situations quickly and intuitively. In the above example of cutting the birthday cake, only the Agent's property <able to perform volitional actions>, the Patient's property <unsturdy>, and the Instrument's property <has a sturdy sharp edge> are necessary for the cutting action. We suggest that it is these discrete, rule-like representations of what is essential for real-world actions that are crucial for flexible comprehension, as they can be applied to any novel combination between actions and entities (Sitnikova, 2003). A given action and its thematic roles (constituting an event) would be understood as long as the perceptual input meets the corresponding minimal requirements.

For instance, imagine that, at a birthday party, a woman starts wriggling a stretch of dental floss across the cake. The observers would probably have little trouble understanding that the woman is cutting the cake, albeit in an unusual way. Now imagine that the woman is wriggling a tissue paper across the cake. This time the observers will probably have no idea what is going on. Nonetheless, the dental floss and tissue paper are both unusual objects at a birthday party and both have semantic properties very different from those of a knife (the object that one expects to be used to cut a cake; e.g., both do not have properties such as <has handle>). We argue that to make such distinctions as between the floss and the tissue paper in the above scenarios, observers would access their knowledge of requirements of the cutting action. The dental floss but not the tissue paper has an edge that is sturdy and sharp enough to cut a relatively unsturdy cake.

It is also possible that comprehenders use action-based requirements to understand the relationships between events. The sequential order of events in most goal-directed activities is not random but is defined by which actions are possible given the current state of environment. For example, in a birthday party scenario, many of the events are possible only if the state of the cake and/or candles matches the requirements for the performed action (e.g., serving a piece of a cake is possible only after it has been cut). These enabling relationships between the events **(p.642)** might be established based on the requirements for each individual action. A similar hypothesis is posed in Chapter 20 in this volume, with the suggestion that observers distinguish between causal, enabling, and preventing relationships between events by calculating how each event influences the current state of the environment in relation to the requirements of the central actions in other events.

In the paragraphs that follow, we first discuss evidence for distinctions between two semantic comprehension systems in the language domain: a system that maps the perceived information on graded semantic representations and a system that utilizes particular semantic requirements of verbs. We then review similar research using static and motion pictures. We suggest that the two mechanisms of language comprehension might be analogous to the systems that use graded semantic representations and action-based requirements to make sense of the visual world.

The experiments that are reviewed in this chapter examine questions of both how comprehenders understand relationships between the elements within individual events and how they understand the relationships between events. Furthermore, throughout the review, we highlight experiments that have used event-related potentials (ERPs). ERPs are electrophysiological brain responses that are recorded via electrode sensors placed on a participant's scalp and are time-locked to the onset of experimental trials of interest (e.g., presentation of target words, object pictures, or visual scenes). They measure brain activity online with a temporal resolution of milliseconds (Cohen, Palti, Cuffin, & Schmid, 1980; Williamson, Kaufman, & Brenner, 1978). This excellent time resolution is valuable in characterizing rapid comprehension processes.

In a typical ERP study, electrophysiological data are collected using 40 to 60 trials per experimental condition and are selectively averaged to obtain a single waveform for each condition. The changes in the neurophysiological activity that give rise to ERPs appear as positive-going or negative-going deflections in the recorded waveform, often referred to as ERP components. These components vary in their distribution across the scalp. Usually, differences in the polarity and/or topography of ERP components between experimental conditions are interpreted as reflecting distinctions in their underlying neuronal sources (Holcomb, Kounios, Anderson, & West, 1999; Kutas, 1993). In the studies described here, such polarity and topography information is used to distinguish **(p.643)** between the neural mechanisms mediating different comprehension processes. In contrast, changes merely in the amplitude or timing of a component across experimental conditions are usually interpreted as indexing modulation of the same neurocognitive process(es) (Holcomb et al., 1999; Kutas, 1993). In the studies described below, the onset, peak latency, and duration of such amplitude changes are used to characterize the time course of the corresponding neurocognitive processes.

Language Comprehension

Graded Semantic Representations in Language Comprehension

Behavioral Studies

In the language domain, it has long been known that in familiar situations, comprehenders tend to fill in information missing from an utterance by using their knowledge of what would normally be expected (Abelson, 1981; Anderson, 1980; Bower et al., 1979; Rumelhart & Ortony, 1977; Schank & Abelson, 1977). For example, imagine a boy talking to his mother: "I've got an invitation to Johnny's birthday party, and I know he really wants a Starfighter Transformer toy. Mom, please!" Clearly, the mother of this boy would understand exactly what he is asking for. This has been argued to be possible because comprehenders store all the likely details of familiar situations in semantic memory within such knowledge structures as schemata (Anderson, 1980; Biederman, Rabinowitz, Glass, & Stacy, 1974; Rumelhart & Ortony, 1977; Zacks & Tversky, 2001; Zacks, Tversky, & Iyer, 2001) or scripts (Abelson, 1981; Bower et al., 1979; Schank & Abelson, 1977). As a result, it is unnecessary to mention all the particulars when people talk about familiar situations; instead it can be safely assumed that listeners will understand the message as long as certain critical components of the situation are mentioned. Experimental evidence for such comprehension by mapping the perceived information on the semantic memory structures comes from, for instance, the reconstructive memory of text (Bower et al., 1979). After reading text passages describing common activities such as eating in a restaurant or visiting a dentist, participants were poor at distinguishing between the statements that were presented in the text and lures that conveyed plausible elements of the described situations but that were not actually presented in the text. Comprehenders also **(p.644)** appeared to tap into their knowledge of the typical temporal order of events in common activities. After reading text passages with sentences presented in a scrambled order, events tended to be recalled in a typical order.

More recent studies have employed a reaction time measure to demonstrate that semantic memory structures are accessed online as comprehenders build up a mental representation of meaning. For example, McRae and colleagues (Ferretti et al., 2001; McRae et al., 2005) demonstrated that online language processing is influenced by common thematic relationships between a given verb and its surrounding noun phrases (NPs). In sentences such as "She was arrested by a cop/crook," expected words such as "cop" were processed faster than unexpected words such as "crook." Note that both types of target words were semantically associated with the target action (e.g., "arresting"); therefore, the processing differences between these conditions suggest that event representations in semantic memory have thematic structure rather than simply tie together related concepts. Three other studies demonstrated that common spatial and temporal relationships between concepts and events are also used in online language comprehension. Richardson, Spivey, Barsalou, and McRae (2003) obtained evidence that verbs access typical spatial properties of their corresponding actions. In their paradigm, participants listened to verbs that commonly refer to vertical actions (e.g., “smash”) or horizontal actions (e.g., “point”). Participants were found to be slower in detecting visually presented stimuli of the corresponding relative to different spatial orientation (e.g., processing “smash” interfered with discriminating targets on the top and bottom of the screen, but not on the left or right). Van der Meer and colleagues showed that verbal stimuli might access knowledge about the common temporal relationships both within and between events. In their studies, pairs of verbal stimuli (including combinations between individual words, word phrases, and sentences) were processed faster when they were presented in the common chronological order than when they were presented in the reversed order (e.g., “shrinking-small” was processed faster than “small-shrinking”—Nuthmann & van der Meer, 2005; “The boy bites off a juicy apple—chew” was processed faster than “The stomach digests the food—swallow”—van der Meer et al., 2002).

An extensive group of studies provides support for the hypothesis that concepts in semantic memory have graded connections, and that **(p.645)** during online language comprehension the perceived information is mapped on these graded representations. Most of these studies used a semantic priming protocol to show that processing time of target words (e.g., “doctor”) decreases from the items preceded by an unrelated prime word (e.g., “cat”) to the items preceded by a moderately related prime (e.g., “accident”), to the items preceded by a strongly related prime (e.g., “nurse”; this is true for the relationships based on feature similarity and association strength—for a review see Hutchison, 2003). An analogous contextual congruency paradigm was also employed for target words embedded in whole sentences. For example, participants might be presented with a sentence stem, “She cleaned the dirt from her,” followed by a target word either predictable (e.g., “shoes”) or acceptable but less probable (e.g., “umbrella”). Comprehenders consistently took longer to respond to unexpected words than to words that were predictable in their preceding context (Fischler & Bloom, 1985; Stanovich & West, 1983). Additional evidence comes from studies examining eye fixations: the duration of readers’ eye fixations tends to be shorter on critical words that are expected relative to those that are unexpected in the preceding context (Morris, 1994; Zola, 1984). Thus, it appears that words can be processed more easily if their corresponding representation is more closely related to the specific field in semantic memory activated by the context.

ERP Studies

More recently, ERP studies have provided further evidence that graded semantic representations are rapidly accessed during language comprehension (within 400 ms after target word onset). These studies recorded ERPs time-locked to the onset of target words in contextual congruency paradigms and identified an electrophysiological marker of the behavioral expectancy effects described above. At approximately 300 ms after word onset, content words evoke a negative-going ERP component that peaks at around 400 ms and accordingly is termed the N400. The magnitude of this N400 is inversely correlated with both the strength of the relationship between the prime and target words in semantic priming paradigms (e.g., Grose-Fifer & Deacon, 2004; Holcomb, 1988, 1993; Kutas & Hillyard, 1989) and the predictability of the target word in the preceding context in sentence paradigms (Kutas & Hillyard, 1980, 1984). For example, in sentences, the final word on trials such as "It was **(p.646)** his first day at *work*" elicits a smaller N400 than the final word on trials such as "He took a drink from the *waterfall*," whereas the final word on trials such as "He took a drink from the *transmitter*" evokes the largest N400. This N400 electrophysiological response has also been reported to critical words that are incongruous with the preceding global context provided by groups of sentences in discourse (van Berkum, Hagoort, & Brown, 1999; van Berkum, Zwitserlood, Hagoort, & Brown, 2003; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Moreover, the modulation of the N400 by congruency is seen both when language is presented visually (reading, e.g., Kutas & Hillyard, 1980) and when language is presented auditorily (speech comprehension, e.g., Holcomb & Neville, 1991a, 1991b). Taken together, these results suggest that the difficulty of mapping the target word on graded semantic representations is reflected by the N400: the closer the relationship between the representation of the eliciting item and the specific semantic memory field activated by the preceding context, the less demanding this mapping process, and the smaller the amplitude of the N400.

Verb-Based Semantic Requirements in Language Comprehension

Even though information stored in graded semantic memory networks can exert powerful effects on language comprehension, these descriptive representations may not be sufficient to achieve accurate comprehension. For example, how would readers arrive at a veridical interpretation of the statement "The humanoid space alien mailed a pencil"? This sentence describes an event that most humans would agree they had never experienced before, and therefore they would not have a prestored representation of this particular event.

Classic linguistic theory posits that, to communicate the relationships between concepts in verbal descriptions, speakers rely on their knowledge of syntactic and semantic requirements governing the correct use of verbs (Fillmore, 1968). Together, these requirements determine which NPs (the verb's arguments) are assigned which thematic roles in a statement that includes a given verb. For example, in a statement involving the verb "mail," two thematic roles have to be considered: the role of an Agent—the entity who is doing the mailing; the role of a Patient—the entity that is being mailed. Syntactically, these roles are expressed by a subject NP and an object NP, respectively. Semantically, the Agent **(p.647)** of the verb "mail" must be able to mail (e.g., be able to perform a volitional action), and its Patient must be "mailable" (e.g., be transportable). Knowing these syntactic and semantic requirements would allow comprehenders to determine the relationships between the verb and its arguments in a sentence. Thus, readers of the statement “The humanoid space alien mailed a pencil” would evaluate each NP against the above requirements so that, for example, the Agent role of the verb “mail” would be assigned to the argument that is a subject NP and is able to mail—to the NP “the humanoid space alien.”

Experimental research provided evidence that such assignment of thematic roles is a rapid online process that heavily depends on syntactic information but may also take into account verb-based semantic requirements (Altmann & Steedman, 1988; Caplan, Hildebrandt, & Waters, 1994; Clifton, 1993; Clifton, Traxler, & Mohamed, 2003; Ferreira & Clifton, 1986; Frazier & Clifton, 1997; Glenberg & Robertson, 2000; Kaschak & Glenberg, 2000; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003; Marslen-Wilson, Brown, & Tyler, 1988; McElree & Griffith, 1995, 1998; Osterhout, Holcomb, & Swinney, 1994). Most importantly, this research suggests that when comprehenders use verb-based semantic requirements to help them in determining relationships between concepts in a sentence, the engaged neurocognitive processes might be different from those that mediate mapping of the input on graded semantic representations.

Behavioral Studies

Behavioral research has established that violations of semantic requirements of verbs are rapidly detected during online language comprehension (Caplan et al., 1994; McElree & Griffith, 1995, 1998). For example, McElree and Griffith (1995) showed that it took readers only a few hundred milliseconds to report such violations in sentences like “Some people *alarm books*” (i.e., books do not have semantic properties necessary for them to become alarmed). In another study, Marslen-Wilson et al. (1988) found that, when monitoring for target words in sentences containing different types of violations, participants were slower to detect words that violated semantic requirements of verbs than to detect nonviolated words. For example, subjects took longer to respond to a target word “guitar” in sentences such as “The crowd was waiting eagerly. John drank the guitar,” than in the control sentences such as “The **(p.648)** crowd was waiting eagerly. John grabbed the guitar.” Interestingly, the time to detect the target words in sentences with verb-based semantic violations was also longer than to detect unexpected target words that did not violate verb-based requirements (e.g., in “The crowd was waiting eagerly. John buried the guitar.”), which suggested that processing verb-based semantic requirements might have a different time course from the process of mapping the target word on graded semantic representations.

The processing distinctions between semantic requirements of verbs and graded semantic representations were more carefully examined in a series of studies by Glenberg and colleagues (Glenberg & Robertson, 2000; Kaschak & Glenberg, 2000). In these experiments, participants’ judgments suggested that they relied on their knowledge of verb-based semantic requirements when asked to comprehend verbally described unusual events.[1] However, the patterns of their responses could not be accounted for by using event representations stored in graded semantic representations. To give an illustration of this line of research, in one of the experiments by Kaschak and Glenberg (2000) participants were asked to read short text passages (e.g., a story about a girl who wanted to prove that she could hit well in baseball; she borrowed a crutch from a person recovering from a twisted ankle and used the crutch to hit an apple). After reading each scenario, participants verified the truth value of a probe statement that was (a) highly relevant for the central action described in the passage (e.g., “the crutch is sturdy”—the crutch sturdiness was necessary for it to be used as a baseball bat), (b) relatively unimportant for the central action (e.g., “the crutch is long”), or (c) irrelevant for the central action but described a scenario that is frequently associated with the critical object (e.g., “the crutch can help with injuries”). Even though the first two probe types were similar in their degree of semantic relatedness to the contextual passages, participants were faster to verify the probes that were highly relevant for the described **(p.649)** central action (sturdy) than the less relevant probes (long). As a measure of semantic relationships between stimuli, most of the experiments by Glenberg and colleagues used Latent Semantic Analysis (Landauer & Dumais, 1997), a computer program that calculates an index of co-occurrence of sets of words in similar contexts. This analysis has been demonstrated to simulate semantic relatedness judgments given by human participants. Interestingly, the highly relevant probes (sturdy) were responded to even faster than the frequent associate probes (injuries), demonstrating that comprehenders did not simply access the most common role played by the target object in the real life but rather used the knowledge of verb-based semantic requirements to integrate these words with the passage context.

ERP Studies

ERP investigations have provided additional insights into how the language comprehension system analyzes verb-based semantic requirements. In our laboratory, we have used the contextual congruency paradigm to compare this mechanism to the processing based on graded semantic representations (Kuperberg, Sitnikova, et al., 2003). We asked participants to read three types of sentences such as the ones below while we recorded ERPs to the target verbs (italicized).

1. Although the young model is completely unaware her fans always *admired . . .*
2. Although the young model is completely unaware her fans always *grieved . . .*
3. Although the young model is completely unaware her hair always *admired . . .*

In sentences like (3), the target verbs are semantically associated with the preceding context[2] (e.g., "admiring" is related to the concept of being a young model who would be expected to have beautiful hair), but their preceding NP argument violates the semantic requirements of these **(p.650)** verbs. In these verb-argument violations, the syntax assigns a thematic role of an Agent around the target verb ("admired") to the preceding subject NP ("her hair"). However, this subject NP does not have semantic properties necessary to perform the described action (e.g., hair can not admire). In contrast, in sentences like (2), the target verb ("grieved") can assign the syntactically defined Agent role to the preceding subject NP ("her fans") because verb-based semantic requirements are not violated (e.g., fans can grieve). In this case, comprehension difficulties arose specifically at the level of relating these sentences to what commonly happens in the real world (e.g., celebrities usually have fans who admire rather than grieve). Our results are shown in Figure 23.1. Confirming prior research, unexpected verbs evoked an increased N400 effect relative to the predictable verbs (Fig. 23.1[*A*]). However, a different pattern of the brain electrophysiological response was evoked by the verb-argument violations. In this condition, the difficulties in integrating the semantically incompatible NPs with their syntactically defined thematic roles were reflected by a later, positive-going ERP wave that started at approximately 500 ms after target verb onset and peaked between 600 and 700 ms (Fig. 23.1[*B*]). Similar results have been also reported by other laboratories (e.g., Hoeks, Stowe, & Doedens, 2004; Kim & Osterhout, 2005; for a review see Kuperberg, 2007).

The late positivity observed in the verb-argument violations was remarkably similar to the P600 ERP component that previously had

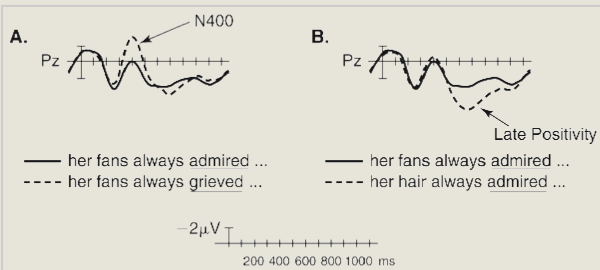

Figure 23.1 . (*A*) Event-related potentials (ERPs) time-locked to unexpected target words compared to ERPs time-locked to predictable target words. (*B*) ERPs time-locked to violations of verb-based semantic requirements compared to ERPs time-locked to predictable target words. Negative voltages are plotted upward.

ERPs shown are at a parietal electrode site.

(p.651) been found for words that could not be easily integrated into the syntactic structure of the preceding sentence (e.g., to syntactic errors—Hagoort & Brown, 2000a; Osterhout & Holcomb, 1992—or when syntactic ambiguity was resolved toward an unpreferred syntactic structure—Osterhout et al., 1994). In fact, in our follow-up study, we directly compared the late positivity to verb-argument violations with the P600 evoked by syntactically anomalous verbs, such as the verb “admires” in (4) below (Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006):

4. Although the young model is completely unaware her fans always *admires . . .*

The results are shown in Figure 23.2, which plots the difference waves obtained by subtracting the ERPs to predictable, syntactically correct target words (e.g., in [1] above) from the ERPs to verb-argument violations (e.g., in [3] above) and syntactic violations (e.g., in [4] above). The two types of violations evoked late positivity effects that were similar in their scalp distribution and timing. Both of these effects were largest at the parietal electrode sites (e.g., Pz) but were less prominent in the more anterior sites (e.g., Fz). Moreover, these effects were similar in their onset (at approximately 500 ms after target verb presentation), peak (at approximately 650 ms), and offset latencies (at approximately 1,000 ms). One interpretation of these similarities might be that the verb-argument violations evoked the late positivity because they were recognized by the processing system as being syntactic, rather than semantic, in nature. Indeed, as discussed above, some evidence suggests that the typical event structure retrieved from graded semantic representations may be used in the online assignment of thematic roles to NPs around the verb (e.g., Ferretti et al., 2001; McRae et al., 2005) and consequently may rapidly influence the syntactic processing of sentences. Perhaps the late positivity is evoked in sentences like (3) above because the subject NP (e.g., hair) is a likely candidate for the Patient role around the target verb (e.g., admired), and as a result these verbs are perceived as morphosyntactic violations (e.g., are perceived as a syntactic error in a sentence “her hair *was* always admired”—this phenomenon has been termed “semantic attraction” of the subject NP to the Patient role—see Kim & Osterhout, 2005).

(p.652)

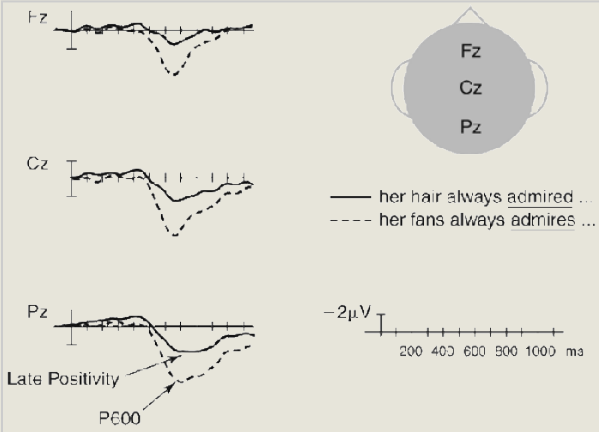

Figure 23.2 . Difference waves obtained by subtracting the event-related potentials (ERPs) time-locked to predictable target words from the ERPs time-locked to violations of verb-based semantic requirements and from the ERPs time-locked to syntactic violations. Shown are waveforms at frontal, central, and parietal electrode sites, whose relative locations on the scalp are indicated on the head diagram (upper right).

An alternative interpretation of the late positivity to verb-argument violations might be that it reflects a semantic integration analysis that is functionally similar to the processing evoked by syntactic anomalies. On this account, it might reflect the process whereby the thematic roles are assigned to the NP arguments by evaluating the *semantic* properties of NPs against the minimal requirements of the target verb (Sitnikova, 2003). Thematic integration between the target verb and its NP arguments, based on at least some of the verb-based semantic requirements, has been suggested to be reflected by a posterior positivity evoked between approximately 200 and 600 ms after the verb presentation (e.g., Bornkessel, Schlesewsky, & Friederici, 2002; Bornkessel, Schlesewsky, & Friederici, 2003). The late positivity evoked by the verb-argument violations may reflect continuing efforts to integrate the target verb with the preceding subject NP by considering whether the properties of the NP match the semantic requirements for some other thematic role around the verb, not the one specified by the syntax. We examined these alternative explanations—based on semantic attraction and processing of verb-based semantic requirements—by recording ERPs to predictable **(p.653)** target words and to two different types of verb-argument violations (Kuperberg, Caplan, et al., 2006). In one sentence type with the violations, such as (3) above, the semantic properties of the subject NP made it a likely candidate for the Patient role around the target verb. In contrast, in the second type of sentences with the violations, such as (5) below, the subject NP does not have the semantic properties required for the Patient role (e.g., it is not possible that “seats would be attended”).

5. Although the lectures are excellent the seats hardly *attend . . .*

If semantic attraction between the subject NP and the Patient thematic role is the main trigger of the late positivity, this effect would be expected in sentences like (3) above but not in sentences like (5). Our results revealed robust late positivity effects to critical verbs in both sentence types (3) and (5) relative to the predictable target verbs, consistent with the hypothesis that this effect reflects attempts to repair the sentences by assigning thematic roles based on verb-based semantic requirements. The obtained ERPs are shown in Figure 23.3. Interestingly, verbs in sentences of type (5) with the Patient-incompatible subject NPs evoked larger late positivities than verbs in sentences of type (3), whose subject NP is compatible with the Patient role. As across these sentence types, semantic association between the target verbs and the preceding context[3] and syntactic complexity are matched, this increase in the positivity effect is likely to reflect an additional mental effort engaged by the attempts to re-assign thematic roles based on verb-based semantic requirements. The late positivity is larger when it **(p.654)** is not possible to find an alternative thematic role for the subject NP around the target verb. Our final follow-up study provided additional evidence that the late positivity to verb-argument violations is not related to the semantic association of the target verb to the context. In this experiment we again recorded ERPs to predictable target words and to two different types of verb-argument violations (Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007). However, this time we compared semantically associated verb-argument violations (some of these sentences were like (3) and some were like (5) above) to verb-argument violations that were not semantically associated with their preceding context. For example, in (6) below, “grieved” is a verb-argument violation (hair can not grieve), and it is semantically unrelated to the “young model and hair” context:

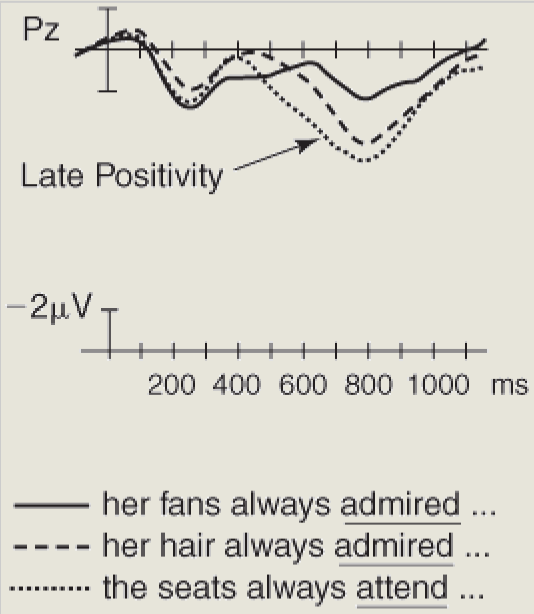

Figure 23.3 . Event-related potentials (ERPs) time-locked to violations of verb-based semantic requirements that were preceded by subject noun phrases either compatible or incompatible with an alternative thematic role of a Patient, compared to ERPs time-locked to predictable target words. Shown are ERPs at a parietal electrode site.

6. Although the young model is completely unaware her hair always *grieved . . .*

As shown in Figure 23.4, the semantically unassociated verb-argument violations evoked a late positivity effect that was even larger than that to the semantically associated verb-argument violations, consistent with the interpretation of the late positivity as reflecting repair attempts based on verb-based semantic requirements.

(p.655)

Taken together, the above ERP findings suggest that during the online sentence processing, comprehenders evaluate NPs against semantic requirements of verbs, and that this analysis is reflected by the late positivity —an ERP component distinct from the N400. This result demonstrates neuroanatomical and temporal distinctions between the processing based on graded semantic representations and verb-based semantic requirements. It is interesting that the late positivity evoked by verb-argument violations is similar to the P600 evoked by syntactic violations, suggesting that these types of anomalies evoke temporally similar processing that is mediated by overlapping neural regions.

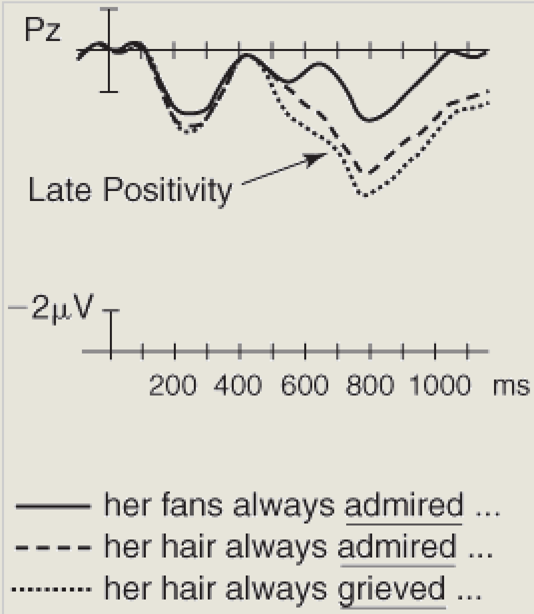

Figure 23.4 . Event-related potentials (ERPs) time-locked to violations of verb-based semantic requirements that either were or were not semantically related to the context, compared to ERPs time-locked to predictable target words. Shown are ERPs at a parietal electrode site.

Language Comprehension: Summary

One line of research on language comprehension has focused on how people rapidly retrieve their knowledge of the world stored within graded representations in semantic memory and use this information as a template for understanding ongoing verbal input in common situations. The other perspective, starting from classic linguistic theory, has been concerned with how comprehenders access their knowledge of semantic requirements stored around verbs and relate this information to NPs within sentences. ERP studies provide evidence that these two streams of cognitive processing may be neurophysiologically distinct. **(p.656)** Mapping the perceived information on graded semantic representations appears to be reflected by the N400 ERP component. Evaluating NPs against verb-based semantic requirements appears to be reflected by the late positivity.

Visual Real-World Comprehension

It is now well established that graded semantic representations are engaged during comprehension not only of language but also other real-world stimuli, including visual images. In contrast, the semantic requirements of verbs have traditionally been assumed to be stored with verbs' lexical entries in the linguistic knowledge system. Even though these requirements may be acquired through learning about the functional requirements of actions in the real world (Pinker, 1989), little is known about whether such requirements play a role in comprehension of real-world events.

Below, we first give an overview of the studies that examined use of graded semantic representations in processing of static pictures. We then examine evidence that, when presented both with the linguistic and visual-world input, comprehenders rapidly integrate properties of the visual environment with verb-based semantic requirements in sentences. Finally, we go on to describe a set of ERP experiments performed in our laboratory that used video clips to examine comprehension of real-world events. These data provide intriguing evidence that to make sense of their visual environment, comprehenders access both graded semantic representations and discrete, rule-like requirements of real-world actions.

Graded Semantic Representations in Visual-World Comprehension

Behavioral Studies

Just as in language, observers of the visual world have been documented to map the perceptual input on their semantic knowledge of common real-world situations. After viewing a scene and then being asked to recount what they saw, viewers frequently incorporate into their accounts expectations of what they think must have been present, even if it was not actually perceived. The phenomenon of boundary extension **(p.657)** (Intraub & Richardson, 1989; Intraub, Bender, & Mangels, 1992; Intraub & Bodamer, 1993; Intraub, Gottesman, Willey, & Zuk, 1996) is one example of such memory distortions. The boundary extension is evidenced by participants' drawings of the previously viewed visual scenes, which often incorporate added elements (e.g., a tree branch over the yard fence that was not present in the original picture). It is also evidenced by recognition tests, in which originally seen visual scenes are often reported as their close-up views and wide-angle foils are frequently reported as old pictures. Similar memory distortions have been reported after viewing video depictions of common sequences of real-life events (e.g., eating at a restaurant; Brewer & Dupree, 1983; Lichtenstein & Brewer, 1980). Viewers are relatively inaccurate in distinguishing the events they have perceived in video clips from plausible foils. The expected order of events in common activities is also known to influence later recall. Even when presented in a scrambled order in videos, the events tend to be later recalled in a usual order (Brewer & Dupree, 1983; Lichtenstein & Brewer, 1980). Importantly, this mapping on semantic memory representations occurs extremely rapidly. The boundary extension effect in memory for visual scenes has been observed even when pictures were presented for only 250 ms at a rate of three stimuli per second (Intraub et al., 1996).

Studies that used semantic priming paradigms with pictures, just as in language, have yielded results suggesting that comprehenders map the perceptual input online on graded representations in their semantic memory. In these studies, pictures that followed a semantically related word or picture were processed faster than pictures that followed semantically unrelated items (Bajo, 1988; Carr, McCauley, Sperber, & Parmelee, 1982; McCauley, Parmelee, Sperber, & Carr, 1980; Sperber, McCauley, Ragain, & Weil, 1979; Theios & Amrhein, 1989). Moreover, the strength of the prime-target connection also influenced the processing times of target pictures (e.g., McEvoy, 1988).

It remains a matter of debate whether pictures and language access the same semantic representations in the brain. According to a *single-code theory* of semantic memory, any stimulus activates common, amodal representations (e.g., Caramazza, Hillis, Rapp, & Romani, 1990; Kroll & Potter, 1984; Pylyshyn, 1980). In contrast, a *multiple-code theory* postulates several forms of semantic knowledge (e.g., visual, verbal, auditory), stored within distinct brain regions and being activated to a **(p.658)** different degree by pictures and words (e.g., Paivio, 1971; 1986; 1991; Shallice, 1988; 1993). Behavioral findings do not clearly support either of these models. For instance, Potter, Kroll, Yachzel, Carpenter, and Sherman (1986) found that plausibility judgments were made just as quickly regarding sentences including just words as sentences in which the final word was replaced with a corresponding picture. For instance, after reading a sentence stem “Paul came to work soaking wet because he forgot his,” participants responded just as fast to the word “umbrella” as to a picture of an umbrella. This was taken to support the single-code theory. In contrast, Paivio (1974) provided evidence that the verbal and image processing mechanisms may be independent, supporting the alternative, multiple-code theory. In this study, participants performed a free recall task after viewing a list of words and pictures. Whereas an immediate repetition of two identical items (either two words or two pictures) in a list produced less than additive effects on free recall, a presentation of an object’s name immediately followed or preceded by the object’s picture resulted in an additive enhancement of recall.

ERP Studies with Static Pictures

ERP findings in a variety of contextual congruency paradigms suggest that comprehenders map visual images on graded semantic representations within approximately 400 ms after stimulus onset. Semantic priming studies have reported smaller N400s to pictures of objects preceded by related compared to unrelated picture primes (Barrett & Rugg, 1990; Holcomb & McPherson, 1994;

McPherson & Holcomb, 1999). Again, the amplitude of this N400 effect was proportional to the relationship strength between the prime and target pictures (e.g., McPherson & Holcomb, 1999). Similarly, object pictures preceded by congruous written sentence contexts evoked a smaller N400 than pictures preceded by incongruous contexts (Federmeier & Kutas, 2001; Ganis, Kutas, & Sereno, 1996). Ganis and Kutas (2003) also showed an N400 effect to individual objects presented in appropriate relative to inappropriate visual scenes. For example, when shown in a “soccer match” background scene, objects such as a soccer ball evoked a smaller N400 than objects such as a toilet paper roll.

Finally, a study from our laboratory showed an N400 effect to final pictures in series of successively presented visual scenes conveying stories (West & Holcomb, 2002). Congruous final scenes (e.g., after being **(p.659)** presented with a series of pictures showing a girl run a race and then fall down, participants viewed the final scene “the girl watching her competitors cross the finish line”) elicited attenuated N400s relative to incongruous final scenes (e.g., “a girl carrying a pot” in the above context).

ERP studies have also addressed the single- versus multiple-code debate about semantic memory: do words and pictures access the same semantic memory representations in the brain? The results generally come out in favor of multiple-code theory: while both linguistic and picture stimuli evoke the N400 component, the distribution of this waveform across the surface of the scalp is different for pictures than for words, suggesting distinct underlying neuronal sources. Whereas the N400 evoked by verbal stimuli is characterized by a parietal-occipital scalp topography (Friederici, Pfeifer, & Hahne, 1993; Hagoort & Brown, 2000b; Holcomb et al., 1999; Kutas & Van Petten, 1994; van Berkum et al., 1999), the negativities elicited by pictures are typically distributed over more anterior electrode sites (Barrett & Rugg, 1990; Hamm, Johnson, & Kirk, 2002; Holcomb & McPherson, 1994; McPherson & Holcomb, 1999; West & Holcomb, 2002). This is illustrated in Figure 23.5, displaying the N400 effects during comprehension of sentences (Kuperberg, Holcomb, Sitnikova, Greve, Dale, & Caplan, 2003) and picture stories (West & Holcomb, 2002). Shown are the difference waves obtained by subtracting the ERPs to predictable target items from the ERPs to unexpected target items. Note that the N400 effect evoked to visual scenes is characterized by a more prolonged time course (beginning at approximately 300 ms after scene onset and lasting until the end of the recording epoch) relative to the N400 effect evoked to written words. One reason for this could be that mapping semantic information about several people and/or objects included within visual scenes on semantic memory representations could unfold over a few hundreds of milliseconds, thus sustaining the enhanced N400 to the incongruous scenario endings. Most importantly, throughout its entire time course, the N400 effect to visual scenes displays more anterior scalp topography than the N400 effect to words. While the verbal effect is maximal at the parietal sites (e.g., Pz), the visual scene effect is primarily evident at the fronto-central sites (e.g., Fz and Cz).

There are also ERP data suggesting that the above neurophysiological distinctions between word and picture stimuli stem from differences in the semantic code that they preferentially access. Within the verbal **(p.660)** domain, the N400 elicited by concrete, easily imageable words (e.g., "dog") is characterized by a more anterior scalp topography than that evoked by abstract words (e.g., "truth"; Holcomb et al., 1999; Kellenbach, Wijers, Hovius, Mulder, & Mulder, 2002; Kounios & Holcomb, 1994; West & Holcomb, 2000). In fact, we have shown that the N400 effect to concrete words (but not abstract words) had a similar anterior topography to the N400 evoked by pictures of individual objects (cf. Holcomb et al., 1999, versus McPherson & Holcomb, 1999). This is demonstrated in Figure 23.6, which includes difference waves (subtractions of ERPs to congruous targets from the ERPs to semantically unrelated targets) obtained for pictures, concrete words, and abstract words. Taken together, these results have been taken to suggest that the N400 component comprises at least two separable negativities that may reflect processing within distinct semantic neural networks (see Holcomb & McPherson, 1994; Holcomb et al., 1999; Kellenbach et al., 2002; McPherson & Holcomb, 1999; Sitnikova, Kuperberg, & Holcomb, 2003; Sitnikova, West, Kuperberg, & Holcomb, 2006; West & Holcomb, 2002). The more anterior negativities elicited by visual images and concrete words might reflect access to graded semantic representations of the visual real world. **(p.661)**

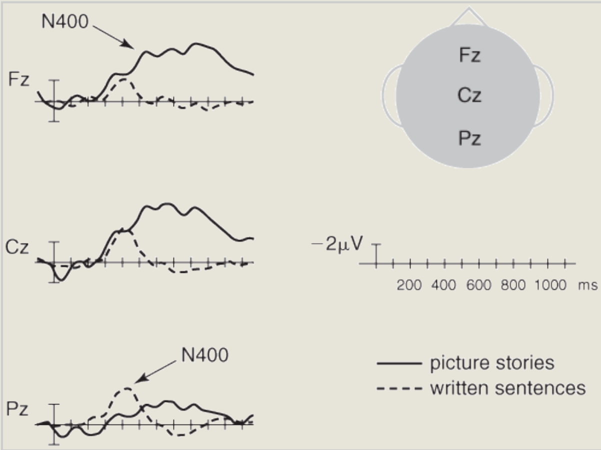

Figure 23.5 . Difference waves obtained by subtracting the event-related potentials (ERPs) time-locked to predictable targets from the ERPs time-locked to unexpected targets in written sentences and picture stories. Shown are waveforms at frontal, central, and parietal electrode sites whose relative locations on the scalp are indicated on the head diagram (upper right).

In contrast, the posterior N400 might reflect activation of the brain regions selectively mediating verbally coded representations.

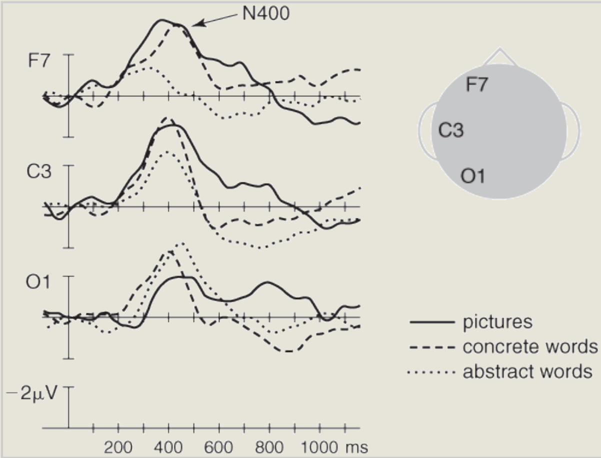

Figure 23.6 . Difference waves obtained by subtracting the event-related potentials (ERPs) time-locked to congruous targets from the ERPs time-locked to semantically unrelated targets, obtained separately for pictures, concrete words, and abstract words. Shown are waveforms at frontal, central, and occipital electrode sites whose relative locations on the scalp are indicated on the head diagram (upper right).

Action-Based Requirements in Visual-World Comprehension: Behavioral Evidence

As discussed in the introduction to this chapter, mapping the perceptual input on graded semantic representations might guide comprehension in familiar situations, such as a traditional birthday party. Such a mechanism, however, is not efficient in less familiar situations and can not explain how people are able to build veridical representations of events that include entities and actions that have not been previously experienced together. We suggest that accurate and flexible comprehension of events in the real world depends on a second semantic mechanism that utilizes discrete, rule-like knowledge of what is necessary for real-world actions.

Several seminal behavioral studies examining interactions between the processing of language and visual environment have established **(p.662)** that object properties perceived in their visual inspection can be rapidly integrated into the online thematic processing of verb-based semantic requirements in sentences (Altmann & Kamide, 1999; Chambers, Tanenhaus, & Eberhard, 2002; Chambers, Tanenhaus, & Magnuson, 2004). For example, Chambers et al. (2004) tracked eye movements of participants who listened to spoken instructions about visual displays of real objects. Thus, participants might be presented with four items: an empty bowl, some flour on a cutting board, and two eggs. One of the eggs would be in a bowl and one in a different container. While viewing this display, participants would hear an instruction: “Pour the egg in the bowl over the flour.” The critical manipulation involved semantic properties of the viewed eggs. In the first, control, condition, both of the eggs were liquid, and hence the display provided no information that would help to disambiguate the object NP “the egg.” In contrast, in the second condition only the egg in the bowl was liquid and therefore could be poured, which disambiguated the NP “the egg.”

This study found that participants used such action affordances, perceived in the visual display, to interpret thematic relationships between constituents in the verbal instructions. In the second but not in the first condition, the propositional phrase “in the bowl” was misinterpreted as a location where the egg had to be poured, as was evident by participants’ anticipatory eye movements toward the empty bowl in the visual display. Similar results were also obtained when the critical variable was the size of objects in visual displays (Chambers et al., 2002).

A similar experiment by Knoeferle, Crocker, Scheepers, and Pickering (2005) suggested that spatial properties of visual-world events also can be combined with the thematic cues in sentences. This study recorded eye movements while participants were presented with static pictures conveying unusual visual events and the corresponding spoken descriptions. Each picture depicted three animate characters involved in two actions; each action involved two characters who performed different roles in the action. For instance, a picture might show a fencer painting a princess who, in turn, was washing a pirate. The verbal descriptions referred to one of the observed actions and were conveyed either by German active sentences (e.g., “Die Prinzessin wird sogleich den Pirat waschen”/“The princess will soon wash the pirate”) or passive sentences (e.g., “Die Prinzessin wird soeben von dem Fechter gemalt”/“The princess is currently painted by the fencer”). The results revealed **(p.663)** that as soon as the linguistic syntactic information that disambiguated the thematic role of the first NP was presented (e.g., “Die Prinzessin” was disambiguated as an Agent or Patient by the temporal adverbs “sogleich/soeben von,” biasing toward the active or passive structure), participants tended to make anticipatory eye movements toward the image of the second character, who was about to be mentioned in the spoken description. Importantly, the only visual information that disambiguated the role of the central character (conveyed by the first NP) in each action was his or her spatial orientation, which made it impossible for the character to play an Agent role in one of the two conveyed actions. Therefore, the above finding indicates that participants were able to use such basic spatial information in assigning the thematic roles to the characters in the pictures.

Taken together, these results suggest, first, that comprehenders are able to rapidly evaluate whether properties of real-world objects meet the semantic requirements of a given verb. This, in turn, leads to a possibility that such processing relies on the knowledge representations analogous to verb-based semantic requirements in language—notably, requirements of real-world actions. Second, the semantic properties of objects appear to be combined with spatial information and possibly also temporal information as comprehenders determine the thematic structure of real-world events.

Based on these findings we put forward a hypothesis for a cognitive mechanism that is able to build veridical representations of real-world events (Sitnikova, 2003). In the visual-world domain, a set of requirements including the semantic properties of entities and the spatiotemporal relationships between them can uniquely constrain specific actions. For example, the cutting action requires that the entity in the Agent role be able to perform cutting (e.g., <have ability for volitional actions>), the entity in the Instrument role have physical properties necessary for cutting (e.g., <have a sturdy sharp edge>), and the entity in the Patient role be cuttable (e.g., <unsturdy>). There are also minimal spatiotemporal requirements for the cutting action (e.g., <the Instrument and the Patient must come in physical contact>). In comprehension of visual events, the correspondence between the perceptual input and the requirements of a given real-world action would allow viewers to identify the event's central action and assign the roles to the involved entities. Of note, employing these discrete, rule-like semantic representations is **(p.664)** fundamentally different from any integration by accessing graded connections between concepts in semantic memory in that this analysis takes into account only a subset of the semantic properties of the visual event—those that are necessary to carry out a given real-world action. As a result, this analysis has great flexibility, as it can be applied to combinations of entities and actions that have not previously been encountered. For example, observers would interpret "wriggling the dental floss across the cake" as "cutting," even if they see such an event for the first time, because the semantic properties of dental floss are consistent with the required properties for cutting a cake. Thus, one interesting possibility is that during visual comprehension a semantic analysis based on the requirements of real-world actions may serve a combinatorial role similar to that played by syntactic processing in language.

ERP Studies in Video Clips: Relationships Between People, Objects, and Actions in Common and Unconventional Visual Real-World Events

Above, we suggested that the comprehension of visual real-world events might involve both mapping on graded semantic representations and the use of discrete, action-based requirements, analogous to the comprehension mechanisms used in language. We have also seen that, in the linguistic domain, ERP data suggest that these two mechanisms might be mediated by the anatomically and temporally distinct neural processes. Difficulties in mapping on graded semantic representations appear to be reflected by the modulation of the N400 waveform. In contrast, difficulties in the thematic analysis based on verb-based semantic requirements appear to be reflected by a somewhat slower late positivity waveform.

In our laboratory, we have recently employed ERPs to determine whether processing based on graded semantic representations and action-based requirements would be neurophysiologically dissociable during visual real-world comprehension (Sitnikova, 2003; Sitnikova et al., 2003; Sitnikova, Holcomb, &

Kuperberg, in press). We reasoned that if these neurocognitive mechanisms are similar to their counterparts in language comprehension, their engagement should evoke similar ERP effects. Mapping of visual events on graded semantic representations would evoke the N400, while evaluating the events against action-based requirements would elicit the late positivity.

(p.665) We explored this hypothesis using naturalistic depictions of real-world events in video clips. Although humans frequently do process static pictures (e.g., in magazines and books), a much more common form of visual comprehension involves the viewing of dynamic images juxtaposed in a continuous flow. Video clips preserve these dynamic properties of the visual environment and are known to evoke perceptual experiences that are remarkably similar to those elicited during comprehending events in the real world (e.g., Levin & Simons, 2000).

In our experiments, we used a contextual congruency paradigm analogous to the one that we previously employed to study semantic processing in the ERP studies of language (Kuperberg, Caplan, et al., 2006; Kuperberg, Kreher, et al., 2007; Kuperberg, Sitnikova, et al., 2003). We produced silent video clips that were about 10 s long. All of these videos were structured in a similar way: a common real-world activity was depicted in a lead-up context and was followed by a congruous or incongruous final scene. ERP recordings were time-locked to the onset of these target scenes.

We used two variations of this contextual congruency paradigm. In the first, we aimed to modulate the difficulty of mapping the final event on graded semantic representations, and in the second we aimed to examine the effects of taxing the analysis based on action-based requirements. In both versions of the paradigm, we used the same set of congruous video clips ending with a predictable final scene. A target object that was introduced in the scenario ending was not seen in the lead-up to these final scenes. For example, in one clip the lead-up context depicted a man squeezing the tire of his bicycle, which appeared soft, and then unscrewing the valve cap on the tire; in the final scene, he used a bike pump (the target object) to fill the tire (Fig. 23.7[*A*]; also see color insert). As described below, the anomalous versions of the scenarios differed depending on whether we aimed to examine the processing based on graded semantic representations or action-based requirements.

Mapping Visual Events on Graded Semantic Representations and the N400>

In our first experiment, we manipulated the predictability of the final scenes in their preceding context (Sitnikova, 2003; Sitnikova et al., in press), extending our earlier work using static pictures described above (West & Holcomb, 2002). The incongruous videos in this experiment **(p.666)** were created by replacing the original context shot in each video with a context shot from another scenario. For example, as outlined above, in the congruous condition, participants would see a man squeeze a tire of his bike, unscrew the tire valve cap, and then use a bike pump to fill the tire (see Fig. 23.7[*A*]; color insert). In the incongruous condition, participants would see a man attempt to open the front door of his house, which turns out to be locked, and fumble in his bag; then the man would use a bike pump to fill a bike tire (see Fig. 23.7[*B*]; color insert). We expected that final scenes in congruous videos would evoke relatively small N400s, as these scenes were predictable and could be easily mapped on the fields in semantic memory activated by the preceding context. In contrast, the incongruous final scenes were expected to evoke increased N400s, as these scenes introduced information that was inconsistent with the semantic representation of the preceding contextual events.

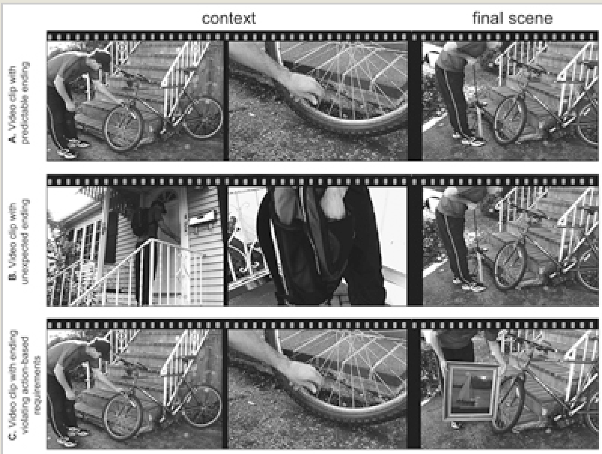

Figure 23.7 . Frames taken from video clips (produced using Canon GL1 digital video camcorder and Adobe digital editing software) used in our contextual congruency paradigm. For each video scenario, shown are two frames illustrating real-world events depicted as a context, followed by a single frame illustrating the final scene. (A) shows a predictable final scene, (B) shows an unexpected final scene, and (C) shows a final scene that is unexpected and violates action-based requirements. The actual video clips may be viewed at http://www.nmr.mgh.harvard.edu/~tatiana/NCMHC.

(p.667) Results of this experiment are shown in Figure 23.8. Starting at approximately 200 ms after their presentation, the final scenes in video depictions of real-world events evoked a robust negative-going ERP that was attenuated by the congruency of the preceding video context. Overall, the morphological, functional, and temporal properties of this effect suggest that it is similar to the N400 previously reported in verbal and static picture paradigms manipulating stimuli predictability (e.g., Ganis et al., 1996; Kutas & Hillyard, 1980; West & Holcomb, 2002). These findings confirm our hypothesis that during comprehension of visual events, the perceptual input is mapped on graded semantic representations. Moreover, they suggest that this mapping starts within a similar time frame as during language comprehension.

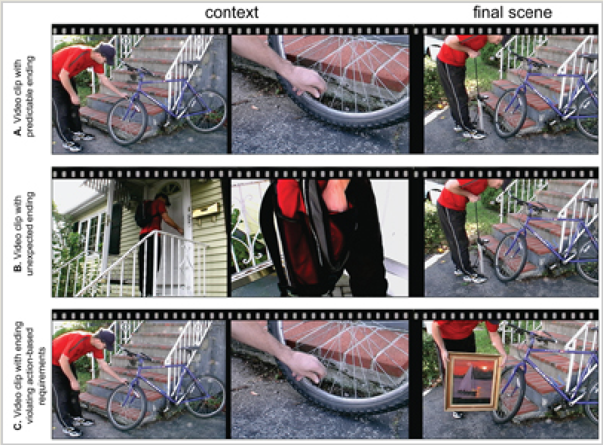

Figure 23.7 . Frames taken from video clips (produced using a Canon GL1 digital video camcorder and Adobe digital editing software) used in our contextual congruency paradigm. For each video scenario, shown are two frames illustrating real-world events depicted as a context, followed by a single frame illustrating the predictable final scene (*A*), the unexpected final scene (*B*), and the final scene that both was unexpected and violated action-based requirements (*C*). The actual video clips may be viewed at http://www.nmr.mgh.harvard.edu/~tatiana/NCMHC.

The time course of the N400 effect evoked by the video stimuli, however, was more prolonged than what is common for the studies with

(p.668) visually presented words (cf. Fig. 23.1[*A*]). The N400 modulation to incongruous relative to congruous final scenes continued for more than 900 ms. This prolonged time course was in keeping with our previous findings with static visual scenes (West & Holcomb, 2002) and could be accounted for by the richness of semantic information conveyed in visual events. Another reason for the prolonged effect could be that presentation of incongruous information in video scenes unfolded continuously over several hundreds of milliseconds, which would lead to the prolongation of the semantic analysis. In line with the latter explanation, similar extended N400s are observed in association with spoken language that also involves processing of critical information over time as the eliciting word unfolds phoneme by phoneme (e.g., Holcomb & Neville, 1991a, 1991b).

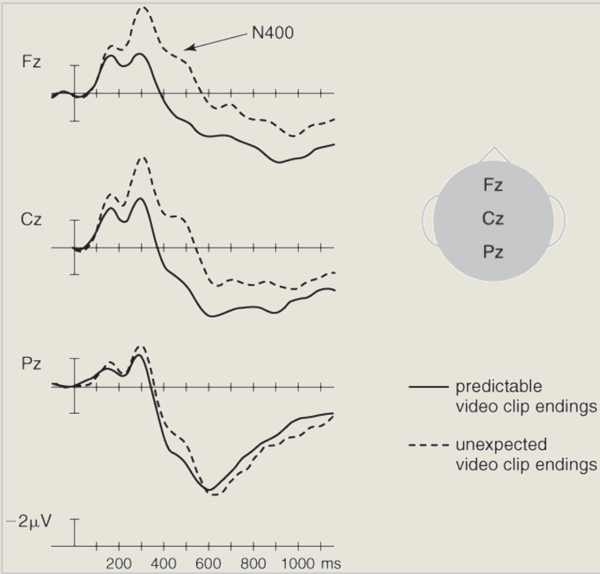

Figure 23.8 . Event-related potentials (ERPs) time-locked to unexpected final scenes compared to ERPs time-locked to predictable final scenes in video scenarios. Shown are waveforms at frontal, central, and parietal electrode sites whose relative locations on the scalp are indicated on the head diagram (upper right).

Another difference between the N400 effects evoked by video scenes and words was in the scalp topography. Just as in our previous studies using static pictures (McPherson & Holcomb, 1999; West & Holcomb, 2002), the N400 to dynamic visual events presented in video clips was evident predominantly over more anterior electrode sites (e.g., Fz electrode site in Fig. 23.8). This finding is consistent with the multiple-code theory of semantic memory: comprehending the visual world and language might preferentially access graded semantic representations supported by distinct neurocognitive systems.

Evaluating Visual Events Against Action-Based Requirements and the Late Positivity

In our next study with video clips, we used the contextual congruency paradigm, which included violations of action-based requirements in addition to manipulations of the final scene's predictability (Sitnikova, 2003; Sitnikova et al., 2003; Sitnikova et al., in press). In this experiment, anomalous scenario endings introduced a single context-inappropriate object that the main character involved in the interaction with another entity in the video clip. For example, after checking the bike tire pressure and unscrewing the tire valve cap, a man twirled a framed painting against the opening of the bike tire valve (see Fig. 23.7[*C*]; color insert). Just as in our previous experiment, these final scenes were unexpected (e.g., twirling a framed painting is not usual when someone is fixing a bike). However, unlike in our previous experiment, these scenes were clearly a continuation of the activity conveyed earlier in the video clip, but did not meet the requirements of the final action constrained by the **(p.669)** preceding context (e.g., twirling a framed painting does not meet the requirements of the action “pumping”). To make any sense of these scenes, participants would need to reevaluate whether the involved entities and their spatiotemporal relationships met the requirements for an alternative action that was also acceptable in the given video scenario (e.g., the framed painting has pointy and sturdy corners: perhaps it is being used to scrape off some relatively soft mud that is stuck to the opening in the tire valve). We anticipated that in addition to evoking an enhanced N400, these target scenes that both were unexpected and violated the requirements of the contextually constrained final action would evoke a large late positivity similar to the one evoked by violations of verb-based semantic requirements in language.

Our results are shown in Figure 23.9.[3] Relative to congruous video endings, the anomalous scenes evoked an increased anterior N400, confirming that it was more difficult to map these unexpected endings on graded semantic representations. In addition, over more posterior scalp regions, these scenes violating action-based requirements evoked an enhanced late positivity. Again, these ERP effects to video materials had a prolonged time course, possibly due to the richness and prolonged presentation of the semantic information. However, their onset latency and overall temporal pattern were similar to the language studies: the N400 started at approximately 250 ms, whereas the late positivity was delayed until approximately 500 ms after target scene presentation. Most importantly, the late positivity effect evoked in video clips resembled in its scalp topography the linguistic late positivity evoked to violations of verb-based semantic requirements—it was widely distributed over more posterior electrode sites, peaking over the parietal sites (a parietal electrode site Pz is shown in Figure 23.9—cf. the late positivity to words in Figures 23.1 and 23.2). These results are interesting as they provide evidence that the processing of action-based requirements during visual-world comprehension might engage neurocognitive mechanisms similar to those mediating the thematic analysis of verb-based semantic requirements in language.

Visual Real-World Comprehension: Summary

Experiments with static and motion pictures suggest that comprehension of the visual world might engage two separate mechanisms, one **(p.670)**

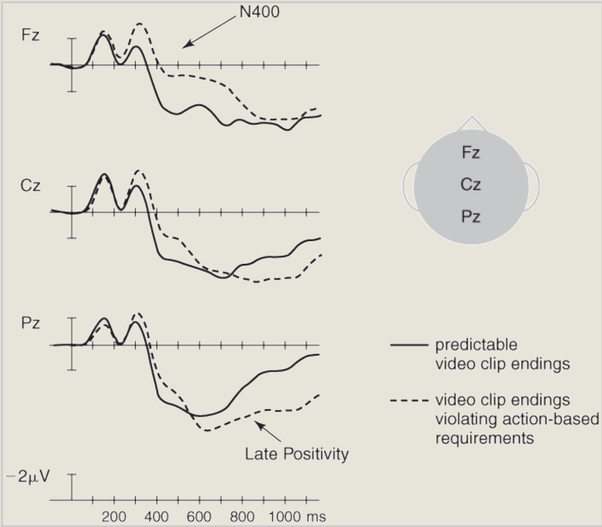

Figure 23.9 . Event-related potentials (ERPs) time-locked to final scenes that were unexpected and violated action-based requirements, compared to ERPs time-locked to predictable final scenes. Shown are waveforms at frontal, central, and parietal electrode sites whose relative locations on the scalp are indicated on the head diagram (upper right).

relying on graded semantic representations and the other relying on the discrete knowledge of action-based requirements. Behavioral studies provide evidence for the mapping of visual-world stimuli on graded semantic representations, and also demonstrate that viewers can rapidly evaluate the visual environment against the semantic requirements of auditorily presented verbs. ERP investigations extend this line of research by revealing neurophysiological distinctions between the two semantic mechanisms during visual real-world comprehension. Several studies with static pictures have demonstrated that contextually unexpected stimuli that are relatively difficult to map on graded semantic representations evoke a large N400. More recently, a set of studies from our laboratory used video depictions of events to dissociate the N400 expectancy effect from the late positivity effect evoked by introducing an additional violation of action-based requirements. Intriguingly, both **(p.671)** of these observed ERP effects to video scenes morphologically, functionally, and temporally resemble the N400 and late positivity evoked in response to semantic violations during language comprehension. The most parsimonious interpretation of this result is that there may be some similarities in the neurocognitive systems involved in language and visual-world comprehension.

Two Semantic Neurocognitive Mechanisms of Comprehension: A Hypothesis and Future Directions

In this chapter we propose a novel hypothesis that attempts to tie together research findings on language and visual-world comprehension. We have outlined prior behavioral and ERP evidence and have presented our own ERP findings suggesting that, in both verbal and visual-world domains, comprehension might be supported by two neuroanatomically and temporally distinct semantic mechanisms. The first mechanism, reflected by the N400 ERP component, appears to continuously access the world knowledge stored in graded semantic representations. The second mechanism, reflected by the late positivity ERP component, appears to access discrete requirements of real-world actions (depicted by verbs in language). There may be a tradeoff between these two types of semantic representation in their utility for integrating the people, objects, and actions during event comprehension, in which the first mechanism is better suited for familiar situations and the second mechanism is better suited for novel situations. In the remainder of this chapter, we consider some of the further questions in this line of research and review some clues to their answers stemming from the currently available data.

One important task for future research is to describe the neural networks mediating each comprehension mechanism. Interestingly, in the nonhuman primate brain, it is well established that there is a functional dichotomy of structures along the anterior–posterior dimension: the posterior cortex appears to specialize in representing sensory information, whereas the frontal lobe is devoted to representing and executing actions. A similar division of labor has been proposed in the human brain (e.g., Fuster, 1997). From this perspective we put forward a question: Is it possible that during event comprehension, graded semantic representations reflecting all prior perceptual experiences and requirements specific to real-world actions are selectively mediated within the **(p.672)** posterior and prefrontal cortices, respectively? To examine this hypothesis, it will be important to use techniques such as functional magnetic resonance imaging and magnetoencephalography that can localize the neural activity in the brain with superior spatial resolution to ERPs.

The currently available neuroimaging data provide evidence for the involvement of primarily the temporal and prefrontal cortices in semantic processing of both language (for a review see Van Petten & Luka, 2006) and visual images (static scenes—e.g., Blondin & Lepage, 2005; video clips—Sitnikova, Coty, Robakis, Holcomb, Kuperberg, & West, 2004). In our laboratory, we have begun the work in the language domain to dissociate between processes based on graded semantic representations and verb-based semantic requirements (Kuperberg, Sitnikova, & Lakshmanan, 2007). Our findings show increased activity in the superior-rostral prefrontal cortex (BA 9) to verb-argument violations (relative to expected verbs) in sentences, suggesting that this brain region may be involved in processing verb-based semantic requirements. Importantly, this activation is not observed to the verbs that are merely unexpected in their preceding sentence context (relative to expected verbs). Other neuroimaging studies have implicated similar superior-rostral prefrontal regions in comprehending relationships between goal-directed and causally related events (Ferstl & von Cramon, 2001, 2002; Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006; Ruby, Sirigu, & Decety, 2002; Tinaz, Schendan, Schon, & Stern, 2006; Xu, Kemeny, Park, Frattali, & Braun, 2005) that, as we argued in the introduction to this chapter, might depend on processing action-based requirements (also see Chapter 22 in this volume, which reviews the role of the prefrontal regions in processing temporal order relationships between real-world events).

The currently available neuropsychological findings also are consistent with the idea that graded semantic representations are mediated within the temporal cortex while action-based requirements are supported within the prefrontal cortex. Patients with temporal lobe damage show severe deficits in word comprehension and picture naming, and the profound temporal damage in semantic dementia tends to result in speech devoid of semantic content while the phonology and grammar are relatively preserved (for a review see Price, 2000). In contrast, patients with frontal lobe damage usually have little difficulty with execution of simple familiar routines but are unable to engage in behaviors that are weakly established and require understanding of a goal and the **(p. 673)** means to achieve it (Shallice & Burgess, 1991). Several studies have suggested that access to graded semantic representations in these patients is relatively spared: they have no difficulty determining which events tend to co-occur in common real-world activities (Sirigu, Zalla, Pillon, Grafman, Agid, & Dubois, 1995, 1996), and they are selectively impaired in sequentially ordering uncommon or unfamiliar rather than routine event sequences (Goel, Grafman, Tajik, Gana, & Danto, 1997; Sirigu et al., 1995). However, these patients with prefrontal damage have difficulties in comprehending relationships between goal-directed and causally related events (Ferstl, Guthke, & von Cramon, 2002; Zalla, Phipps, & Grafman, 2002) and ordering events in goal-directed sequences (Sirigu, Cohen, Zalla, Pradat-Diehl, Van Eeckhout, Grafman, & Agid, 1998).

The second important task for future studies is to delineate functional differences in neural processing between the two comprehension systems. Some insights with regard to the neural mechanisms able to support graded semantic representations and discrete, rule-like, action-based requirements come from research in computational neuroscience. In connectionist networks, acquisition and use of graded semantic knowledge have been simulated by means of variation in synaptic weights representing connection strengths (based on feature similarity or association strength) between the learned concepts (for a review see Hutchison, 2003). However, more recent studies have demonstrated that the neurobiological mechanisms specific to the prefrontal cortex (which support updating of active maintenance contingent on the presence of a reward) can lead to self-organization of discrete, rule-like representations coded by patterns of activity (e.g., distinct sets of units with high synaptic weights—e.g., Rougier, Noelle, Braver, Cohen, & O’Reilly, 2005). This computational model supports adaptive processing in novel situations by searching for the appropriate pattern of activity, which eliminates the need to learn a new set of connection strengths. Even though so far this model has only been shown to simulate human participants’ performance on specific rule-driven behavioral tasks designed for experimental settings (e.g., the Stroop task), it is possible that the neural mechanisms within the prefrontal cortex can also support patterns of activity coding requirements of real-world actions. Specifically, through breadth of learning experience with actions that achieved or failed to achieve their goal (i.e., either resulted in a "reward" or not), these prefrontal mechanisms can identify the pattern of activity present **(p.674)** across all instances of achieving a specific goal.[4] In comprehension, searching for such a pattern of activity, representing the parameters necessary for a given action, would allow flexibility in recognizing actions in novel circumstances.

Based on this model of coding action-based requirements as patterns of neuronal activity, we can make a specific prediction for future research. As a given pattern of neuronal activity only represents the properties necessary for a given action, this mechanism should be insensitive to event properties that are not necessary for the perceived action. For example, in the "cutting the cake" scenario, very different objects such as a plate, a tape measure, or dental floss, when they are used as Instruments of cutting, should access the same pattern of neuronal activity coding the requirement <the Instrument of the cutting action must have a sturdy sharp edge>. Interestingly, electrophysiological studies in nonhuman primates have already obtained some evidence that prefrontal neurons display such a discrete pattern of response to categories of visual stimuli (Freedman, Riesenhuber, Poggio, & Miller, 2001, 2002, 2003) and match-mismatch relationships (Wallis, Anderson, & Miller, 2001; Wallis & Miller, 2003) that are defined by their functional relevance (for a review see Miller, Freedman, & Wallis, 2002; Miller, Nieder, Freedman, & Wallis, 2003).

Yet another research task will be to determine how the two neurocognitive mechanisms of comprehension interact with each other in real time. Rapid reciprocal influences between these processing streams can be mediated by cortico-cortical axons interlinking cortical regions (e.g., Fuster, 1997). Recent ERP data in language comprehension give some evidence for these rapid interactions: outputs of one stream appear to immediately influence processing within the other (for review see Kuperberg, 2007). For example, detection of the violated verb-based semantic requirements (e.g., in "Although the young model is completely unaware her hair always *grieved* . . ."—Kuperberg, Kreher, et al., 2007) seems to reduce attempts to map the target word on graded semantic representations—words like "grieved" that are semantically **(p.675)** unrelated to the context did not evoke a large N400. On the other hand, to motivate an extended processing based on verb-based semantic requirements, it might be necessary that the sentence activate a specific field in graded semantic representations. In sentences with less constraining context, the verb-argument violations (e.g., in "The library books had been *regretting* . . .") evoked an N400 effect rather than the late positivity (Kim & Osterhout, 2005). Given the functional similarities between the semantic integration processes between language and visual event comprehension that we discussed above, it is likely that similar dynamics might take also place during visual real-world comprehension.

Conclusion

In this chapter, we suggest that comprehenders in the real world engage two distinct neurocognitive mechanisms that may be analogous to those employed during language processing. One system might be adapted for efficiency in everyday predictable life. Another might be fine-tuned for flexible ability to rapidly make sense of even novel or unusual situations. Although most of the existing evidence for this perspective is in the language domain, we have presented recent electrophysiological data suggesting that this framework might generalize to real-world comprehension. Future experiments will need to test this theoretical perspective.

References

Bibliography references:

Abelson, R. P. (1981). Psychological status of the script concept. *American Psychologist*, *36*, 715–729.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*(3), 191–238.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

Anderson, J. R. (1980). Concepts, propositions, and schemata: What are the cognitive units? *Nebraska Symposium on Motivation*, *28*, 121–162.

Bajo, M. T. (1988). Semantic facilitation with pictures and words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*(4), 579–589.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain Cognition*, *14*(2), 201–212.

Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W., Jr. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, *103*(3), 597–600.

(p.676) Blondin, F., & Lepage, M. (2005). Decrease and increase in brain activity during visual perceptual priming: An fMRI study on similar but perceptually different complex visual scenes. *Neuropsychologia*, *43*(13), 1887–1900.

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2002). Beyond syntax: language-related positivities reflect the revision of hierarchies. *Neuroreport*, *13*(3), 361–364.

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2003). Eliciting thematic reanalysis effects: The role of syntax-independent information during parsing. *Language & Cognitive Processes*, *18*(3), 269–298.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*(2), 177–220.

Brewer, W. F., & Dupree, D. A. (1983). Use of plan schemata in the recall and recognition of goal-directed actions. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *9*(1), 117–129.

Caplan, D., Hildebrandt, N., & Waters, G. S. (1994). Interaction of verb selectional restrictions, noun animacy and syntactic form in sentence processing. *Language & Cognitive Processes*, *9*(4), 549–585.

Caramazza, A., Hillis, A. E., Rapp, B. C., & Romani, C. (1990). The multiple semantics hypothesis: Multiple confusions? *Cognitive Neuropsychology*, *7*(3), 161–189.

Carr, T. H., McCauley, C., Sperber, R. D., & Parmelee, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception & Performance*, *8*(6), 757–777.

Chambers, C. G., Tanenhaus, M. K., & Eberhard, K. M. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory & Language*, *47*(1), 30–49.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*(3), 687–696.

Clifton, C., Jr. (1993). Thematic roles in sentence parsing. *Canadian Journal of Experimental Psychology*, *47*(2), 222–246.

Clifton, C. J., Traxler, M. J., & Mohamed, M. T. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory & Language*, *49*(3), 317–334.

Cohen, D., Palti, Y., Cuffin, B. N., & Schmid, S. J. (1980). Magnetic fields produced by steady currents in the body. *Proceedings of the National Academy of Sciences of the USA*, *77*(3), 1447–1451.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *27*(1), 202–224.

Ferreira, F., & Clifton, J., C. (1986). The independence of syntactic processing. *Journal of Memory & Language*, *25*, 348–368.

Ferretti, T., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory & Language*, *44*, 516–547.

(p.677) Ferstl, E. C., Guthke, T., & von Cramon, D. Y. (2002). Text comprehension after brain injury: Left prefrontal lesions affect inference processes. *Neuropsychology*, *16*(3), 292–308.

Ferstl, E. C., & von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension: An event- related fMRI study. *Brain Research Cognitive Brain Research*, *11*(3), 325–340.

Ferstl, E. C., & von Cramon, D. Y. (2002). What does the frontomedian cortex contribute to language processing: Coherence or theory of mind? *Neuroimage*, *17*(3), 1599–1612.

Fillmore, C. (1968). The case for Case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Reinhart and Winston.

Fischler, I. S., & Bloom, P. A. (1985). Effects of constraint and validity of sentence contexts on lexical decisions. *Memory & Cognition*, *13*(2), 128–139.

Frazier, L., & Clifton, C., Jr. (1997). Construal: Overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, *26*(3), 277–295.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, *291*(5502), 312–316.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2002). Visual categorization and the primate prefrontal cortex: Neurophysiology and behavior. *Journal of Neurophysiology*, *88*(2), 929–941.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, *23*(12), 5235–5246.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Brain Research Cognitive Brain Research*, *1*(3), 183–192.

Fuster, J. M. (1997). Network memory. *Trends in Neuroscience*, *20*(10), 451–459.

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Brain Research Cognitive Brain Research*, *16*(2), 123–144.

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, *8*(2), 89–106.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory & Language*, *43*(3), 379–401.

Goel, V., Grafman, J., Tajik, J., Gana, S., & Danto, D. (1997). A study of the performance of patients with frontal lobe lesions in a financial planning task. *Brain*, *120*(Pt 10), 1805–1822.

Grose-Fifer, J., & Deacon, D. (2004). Priming by natural category membership in the left and right cerebral hemispheres. *Neuropsychologia*, *42*(14), 1948–1960.

Hagoort, P., & Brown, C. M. (2000a). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, *38*(11), 1531–1549.

Hagoort, P., & Brown, C. M. (2000b). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia*, *38*(11), 1518–1530.

(p.678) Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, *113*(8), 1339–1350.

Hoeks, J. C., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Brain Research: Cognitive Brain Research*, *19*(1), 59–73.

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain Language*, *35*(1), 66–85.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*(1), 47–61.

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *25*(3), 721–742.

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain Cognition*, *24*(2), 259–276.

Holcomb, P. J., & Neville, H. J. (1991a). The electrophysiology of spoken sentence processing. *Psychobiology*, *19*, 286–300.

Holcomb, P. J., & Neville, H. J. (1991b). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology*, *19*(4), 286–300.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomics Bulletin Review*, *10*(4), 785–813.

Intraub, H., Bender, R. S., & Mangels, J. A. (1992). Looking at pictures but remembering scenes. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *18*(1), 180–191.

Intraub, H., & Bodamer, J. L. (1993). Boundary extension: Fundamental aspect of pictorial representation or encoding artifact? *Journal of Experimental Psychology: Learning, Memory & Cognition*, *19*(6), 1387–1397.

Intraub, H., Gottesman, C. V., Willey, E. V., & Zuk, I. J. (1996). Boundary extension for briefly glimpsed photographs: Do common perceptual processes result in unexpected memory distortions? *Journal of Memory & Language*, *35*(2), 118–134.

Intraub, H., & Richardson, M. (1989). Wide-angle memories of close-up scenes. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *15*(2), 179–187.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory & Language*, *43*(3), 508–529.

Kellenbach, M. L., Wijers, A. A., Hovius, M., Mulder, J., & Mulder, G. (2002). Neural differentiation of lexico-syntactic categories or semantic features? Event-related potential evidence for both. *Journal of Cognitive Neuroscience*, *14*(4), 561–577.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory & Language*, *52*(2), 205–225.

(p.679) Klix, F. (2001). The evolution of cognition. *Journal of Structural Learning and Intelligence Systems*, *14*, 415–431.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95–127.

Knutson, K. M., Wood, J. N., & Grafman, J. (2004). Brain activation in processing temporal sequence: An fMRI study. *Neuroimage*, *23*(4), 1299–1307.

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *20*(4), 804–823.

Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior*, *23*, 39–66.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49.

Kuperberg, G. R., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. J. (2006). Neural correlates of processing syntactic, thematic and semantic relationships in sentences. *Language and Cognitive Processes*, *21*, 489–530.

Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, *15*(2), 272–293.

Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, *100*(3), 223–237.

Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. M., & Holcomb, P. J. (2006). Making sense of discourse: An fMRI study of causal inferencing across sentences. *NeuroImage*, *33*(1), 343–361.

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Brain Research Cognitive Brain Research*, *17*(1), 117–129.

Kuperberg, G. R., Sitnikova, T., & Lakshmanan, B. (2007). Semantic violations of action and morphosyntactic agreement violations recruit an overlapping neural network: Evidence from functional magnetic resonance imaging (submitted).

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language & Cognitive Processes*, *8*(4), 533–572.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163.

Kutas, M., & Hillyard, S. A. (1989). An electrophysiological probe of incidental semantic association. *Journal of Cognitive Neuroscience*, *1*, 38–49.

Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego, CA: Academic Press.

(p.680) Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's Problem: The latent semnatic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Levin, D. T., & Simons, D. J. (2000). Perceiving stability in a changing world: Combining shots and integrating views in motion pictures and the real world. *Media Psychology*, *2*, 357–380.

Lichtenstein, E. D., & Brewer, W. F. (1980). Memory for goal-directed events. *Cognitive Psychology*, *12*, 412–445.

Marslen-Wilson, W., Brown, C. M., & Tyler, L. K. (1988). Lexical representations in spoken language comprehension. *Language & Cognitive Processes*, *3*(1), 1–16.

McCauley, C., Parmelee, C. M., Sperber, R. D., & Carr, T. H. (1980). Early extraction of meaning from pictures and its relation to conscious identification. *Journal of Experimental Psychology: Human Perception Performance*, *6*(2), 265–276.

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*(1), 134–157.

McElree, B., & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*(2), 432–460.

McEvoy, C. L. (1988). Automatic and strategic processes in picture naming. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *14*(4), 618–626.

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, *36*(1), 53–65.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition*, *33*(7), 1174–1184.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.

Miller, E. K., Freedman, D. J., & Wallis, J. D. (2002). The prefrontal cortex: Categories, concepts and cognition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *357*(1424), 1123–1136.

Miller, E. K., Nieder, A., Freedman, D. J., & Wallis, J. D. (2003). Neural correlates of categories and concepts. *Current Opinion Neurobiology*, *13*(2), 198–203.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *20*(1), 92–103.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading*. Hillsdale, NJ: Laurence Erlbaum Associates.

Nowak, M. A., Plotkin, J. B., & Jansen, V. A. (2000). The evolution of syntactic communication. *Nature*, *404*(6777), 495–498.

Nuthmann, A., & van der Meer, E. (2005). Time’s arrow and pupillary response. *Psychophysiology*, *42*(3), 306–317.

(p.681) Osterhout, L., & Holcomb, P. J. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory & Language*, *31*, 785–806.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *20*(4), 786–803.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.

Paivio, A. (1974). Spacing of repetitions in the incidental and intentional free recall of pictures and words. *Journal of Verbal Learning & Verbal Behavior*, *13*, 497–511.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General*, *115*(3), 281–294.

Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy*, *197*(Pt 3), 335–359.

Pylyshyn, Z. W. (1980). Computation and cognition: Issues in the foundations of cognitive science. *Behavioral and Brain Sciences*, *3*, 111–132.

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

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences of the USA*, *102*(20), 7338–7343.

Ruby, P., Sirigu, A., & Decety, J. (2002). Distinct areas in parietal cortex involved in long-term and short-term action planning: A PET investigation. *Cortex*, *38*(3), 321–339.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 97–135). Hillsdale, NJ: Erlbaum.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Shallice, T. (1988). Specialisation within the semantic system. *Cognitive Neuropsychology*, *5*(1), 133–142.

Shallice, T. (1993). Multiple semantics: Whose confusions? *Cognitive Neuropsychology*, *10*(3), 251–261.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*(Pt 2), 727–741.

(p.682) Sirigu, A., Cohen, L., Zalla, T., Pradat-Diehl, P., Van Eeckhout, P., Grafman, J., et al. (1998). Distinct frontal regions for processing sentence syntax and story grammar. *Cortex*, *34*(5), 771–778.

Sirigu, A., Zalla, T., Pillon, B., Grafman, J., Agid, Y., & Dubois, B. (1995). Selective impairments in managerial knowledge following pre-frontal cortex damage. *Cortex*, *31*(2), 301–316.

Sirigu, A., Zalla, T., Pillon, B., Grafman, J., Agid, Y., & Dubois, B. (1996). Encoding of sequence and boundaries of scripts following prefrontal lesions. *Cortex*, *32*(2), 297–310.

Sitnikova, T. (2003). *Comprehension of videos of real-world events: Electrophysiological evidence*. Unpublished doctoral dissertation. Medford, MA: Tufts University.

Sitnikova, T., Coty, A., Robakis, D., Holcomb, P. J., Kuperberg, G. R., & West, W. C. (2004). FMRI correlates of comprehending real-world events. *Neuroscience 2004, the Society for Neuroscience's 34th Annual Meeting*. *Society for Neuroscience Abstracts*.

Sitnikova, T., Holcomb, P. J., & Kuperberg, G. (in press). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*.

Sitnikova, T., Kuperberg, G., & Holcomb, P. J. (2003). Semantic integration in videos of real-world events: An electrophysiological investigation. *Psychophysiology*, *40*(1), 160–164.

Sitnikova, T., West, W. C., Kuperberg, G. R., & Holcomb, P. J. (2006). The neural organization of semantic memory: Electrophysiological activity suggests feature-based segregation. *Biological Psychology*, *71*(3), 326–340.

Sperber, R. D., McCauley, C., Ragain, R. D., & Weil, C. M. (1979). Semantic priming effects on picture and word processing. *Memory and Cognition*, *7*(5), 339–345.

Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, *112*(1), 1–36.

Theios, J., & Amrhein, P. C. (1989). Theoretical analysis of the cognitive processing of lexical and pictorial stimuli: Reading, naming, and visual and conceptual comparisons. *Psychological Review*, *96*(1), 5–24.

Tinaz, S., Schendan, H. E., Schon, K., & Stern, C. E. (2006). Evidence for the importance of basal ganglia output nuclei in semantic event sequencing: An fMRI study. *Brain Research*, *1067*(1), 239–249.

Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *31*(3), 443–467.

van Berkum, J. J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657–671.

van Berkum, J. J., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Brain Research: Cognitive Brain Research*, *17*(3), 701–718.

(p.683) van der Meer, E., Beyer, R., Heinze, B., & Badel, I. (2002). Temporal order relations in language comprehension. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *28*(4), 770–779.

Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain Language*, *97*(3), 279–293.

Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, *411*(6840), 953–956.

Wallis, J. D., & Miller, E. K. (2003). From rule to response: Neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology*, *90*(3), 1790–1806.

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience*, *12*(6), 1024–1037.

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Brain Research: Cognitive Brain Research*, *13*(3), 363–375.

Williamson, S. J., Kaufman, L., & Brenner, D. (1978). Latency of the neuromagnetic response of the human visual cortex. *Vision Research*, *18*(1), 107–110.

Wood, J. N., & Grafman, J. (2003). Human prefrontal cortex: Processing and representational perspectives. *Nature Reviews Neuroscience*, *4*(2), 139–147.

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *Neuroimage*, *25*(3), 1002–1015.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, *130*(1), 29–58.

Zalla, T., Phipps, M., & Grafman, J. (2002). Story processing in patients with damage to the prefrontal cortex. *Cortex*, *38*(2), 215–231.

Zola, D. (1984). Redundancy and word perception during reading. *Perceptual Psychophysiology*, *36*(3), 277–284. **(p.684)**

Notes:

(1) . In some of these experiments the actions were not described by a single verb; rather, their meaning became clear from the overall context—in fact, similar effects were found both with conventional verbs (e.g., “to hit”) and novel verbs (e.g., “to crutch”). As a result, this data provides some evidence that not only semantic constraints stored in verbs’ memory representations in the language system but also requirements of real-world actions are used in language comprehension.

(2) . A separate rating study has shown that these target verbs were semantically associated with both the preceding subject NPs and the other content words in the preceding context (see Kuperberg, Caplan, et al., 2006; Kuperberg, Sitnikova, et al., 2003).

(3) . This experiment was carried out using *two types of video stimuli*: in one study, the video clips were continuous with no cuts between scenes (Sitnikova et al., 2003); in the other study, the final scene was shown after a cut (Sitnikova, 2003; Sitnikova et al., in press). This experiment was also carried out using *two different task instructions*: in one study, participants were asked to explicitly decide whether the presented sequence of events would commonly be witnessed in everyday life or not (Sitnikova et al., 2003). In contrast, in the other study participants were not instructed to classify the scenarios and instead answered occasional questions about their content (Sitnikova, 2003). These manipulations helped us to demonstrate that cuts in video clips or performing an additional classification task do not disrupt naturalistic comprehension. Shown are the ERPs evoked in video clips with cuts while participants performed the scenario classification task.

(4) . It is also possible that the same prefrontal mechanisms can code requirements of mental activities such as “admiring” that might be learned through feedback (from other individuals) about correct and incorrect attributing of such mental activities to oneself and surrounding entities.