Higher level language processes in the brain: Inference and Comprehension Processes

International Conference at the Hanse-Advanced Study Institute/Germany
June, 21-25, 2003

Abstracts
FMRI signal when people generate causal inferences during stories.

Mark J. Beeman
Northwestern University

Although the majority of psycholinguistic research concerns how people process words or sentences, the bulk of human communication occurs at the level of discourse. Some language comprehension processes may be far more evident as people comprehend complex discourse than when they comprehend words or sentences. For instance, when people comprehend stories, they generate causal inferences as necessary to fill coherence gaps. Studies with neurological patients and with normal participants demonstrating inference priming of lateralized target words suggest that multiple, bilateral, components of semantic processing are necessary for optimally drawing coherence inferences. I will discuss evidence of such processing from several methodologies (brain damage, visual field studies with healthy subjects, and neuroimaging) and types of stimuli (multi-word prime-target pairs, stories that promote inferences, and insight-like verbal problems). Results are interpreted in a framework positing bilateral activation, integration, and selection of semantic representations, each hemisphere performing slightly different computations within each semantic process. Specifically, I will argue that the left hemisphere performs relatively finer coding of semantic information, which is extremely efficient for most straightforward language comprehension; whereas the right hemisphere performs relatively coarser coding of semantic information, which is especially useful when people understand certain discourse constructions and when semantic integration is difficult, and adds richness to discourse representations generally.
Inference by Coincidence and the Extraction of Propositional Information from Text

Simon Dennis
University of Colorado

The ability to extract propositional information from text is a fundamental prerequisite for models of sentence processing and text comprehension. That people encode sentences in something like propositions, and not simply as lists of words, is well established (see Kintsch 1998 for a summary of the evidence). However, the exact nature of these representations and how the cognitive system goes about the task of transforming a string of words into these propositional representations remains an open question. Current systems require substantive designer intervention to build grammars and conceptual schemas that specify the way in which information is structured (e.g. Lenat 1995). However, as yet no generic account of how these templates could be acquired by people has emerged.

The Syntagmatic Paradigmatic (SP) model offers a solution to this problem. Like approaches such as Latent Semantic Analysis (LSA, Landauer & Dumais 1997), the SP model is completely data driven. Rather than propose pre-existing syntactic and conceptual rules, the SP model presumes that people store large numbers of sentence traces. Processing a sentence then becomes a matter of retrieving appropriate traces from memory and aligning them with the target sentence. To provide precise definitions of retrieval and alignment the model introduces a Bayesian interpretation of String Edit Theory and incorporates a version of the Expectation Maximization algorithm to learn edit probabilities from text corpora.

Unlike LSA, however, the SP model captures propositional knowledge. To demonstrate this property, the model was applied to answering questions about who won tennis matches given un-annotated news articles from the Association of Tennis Professionals (ATP) website. Note that discerning the winner of a match requires that the model be able to assign the winner role to the appropriate constituents of a sentence and cannot be achieved reliably from word level properties. Furthermore, automatic inference occurs as an emergent property of the mechanism and in the demonstration was found to account for a larger proportion of the performance of the model than literal retrieval. This "inference by coincidence" may provide an explanation for the quick, reliable and ubiquitous inferencing abilities that people display.
Evidence is accumulating that readers routinely activate perceptual symbols during language comprehension (Kaschak & Glenberg, 2000; Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002; Zwaan & Yaxley, in press). Perceptual symbols are modal representations that have an analog relationship with their referents because of their correspondence to the perceptual state that produced it (Barsalou, 1999).

Zwaan et al. (2002) showed that response times were slower for a yes/no decision as to whether a picture occurred in a preceding sentence when this picture mismatched the implied shape of an object in the preceding sentence, the Mismatch Condition, than when the picture matched the implied shape of an object in that sentence, the Match Condition. In the Mismatch condition, the sentence was for example: “The ranger saw an eagle in the sky.” The sentence was followed by a line drawing of an object, an eagle with its wings drawn in, which mismatched the implied shape of the eagle in the sentence. In the Match condition, the picture of the line drawing of the object would have been an eagle with its wings outstretched, or the sentence would have been “The ranger saw an eagle in its nest.” The results suggest that the representation of an eagle in the sky with its wings outstretched is captured by a perceptual representation of the sentence, but not by the argument EAGLE in a propositional network (van Dijk & Kintsch, 1983).

Analog representations of objects in language comprehension contribute to the construction of situation models. Situation models are mental representations of the state of affairs described in a text (van Dijk & Kintsch, 1983) and are typically viewed as networks of amodal propositions. Successful comprehension of a text occurs when readers establish representations of the text at the surface level, propositional level, and situation model level of the text (van Dijk, & Kintsch, 1983). If perceptual symbols are routinely used during comprehension of text, then a purely propositional representation of the surface structure, idea units, and situation model may not be able to capture the richness of visual representation of objects in the text, whereas an analog representation of the text can.

The aim of the present study was to assess whether older adults would display an analog representation of information in language similar to what has previously been shown by younger adults (Stanfield & Zwaan, 2001; Zwaan et al, 2002). Because research on situation models and aging showed similar construction and updating processes for younger and older adults (Radvansky, 1999), we expected older adults to show the mismatch effect in a similar manner as the younger adults, or perhaps even show a larger mismatch effect. In the latter case, older adults’ relatively stronger situation model would be more difficult to override with inconsistent information contained in the picture compared to the situation model of younger adults.

Nineteen younger (mean age 18.25, SD=.64) and 19 older adults (mean age 74.4, SD=7.39) participated in the experiment. Materials included 84 line drawings, 56 of which served as filler items. The remaining 28 experimental items consisted of object pairs with each member of the pair denoting a different shape of the object.
Eighty-four sentences were created to accompany the pictures; 56 filler sentences and 28 experimental sentences (14 sentences implying one shape of a given object, and 14 sentences implying the alternate shape). The filler sentences all mentioned an object (by way of a concrete noun) other than the subsequently pictured object, and thus required a “no” response on the recognition task. The experiment was run on a PC using the E-prime 1.0 software. Responses were recorded via the keyboard, using the “j” for yes and “f” for no responses.

A 2 (match/mismatch) by 2 (age group) analysis of variance (ANOVA) was conducted on response latencies and accuracy showing a significant mismatch effect: the responses were faster when sentence and picture matched compared to when they mismatched, an effect of age, with younger adults having shorter response latencies than older adults, and a condition by age group interaction with the mismatch effect being larger for the older adults than for the younger adults (mean match young=809ms, SD=146.5, mean mismatch young=864, SD=173.5; mean match old =1264, SD=264.0, mean mismatch old=1407, SD=290.7). This interaction remained significant after a z score transformation (described in Faust et al., 1999) that took the greater variability in response times of older adults into account. The participant and item analyses on accuracy yielded a main effect of age group only indicating greater accuracy for younger than older adults.

The present results suggest that, like younger adults (Zwaan et al., 2002), older adults use perceptual information in sentence comprehension. Our findings are consistent with other studies showing similar performance in the construction of situation models (Radvansky, 1999). The results also suggest that older adults use perceptual information to a greater extent than younger adults, even when response time is controlled for age-related slowing.

This research was supported by grant MH 63972 from the National Institute on Mental Health.

References
Higher level language processes have only recently become the subject of neuroimaging studies (for reviews see Bookheimer, 2002; Gernsbacher & Kaschak, in press). In addition to the “classical” perisylvian language areas in the dominant hemisphere, a number of brain areas have been shown to be engaged during text comprehension. For instance, the left fronto-median cortex has been shown to be important for coherence processes (Ferstl & von Cramon, 2001; Ferstl, Guthke, & von Cramon, 2002), and the right hemisphere has been implicated in a number of neuroimaging studies on semantic, contextual and pragmatic aspects of language. However, due to the paucity of available data, it has not yet been possible to resolve some inconsistencies and to convincingly map these different regions onto dissociable processes postulated in psycholinguistic theories of text comprehension.

In the present talk, we focus on the process of situation model building (van Dijk & Kintsch, 1983). This mental model of the global content is an integration of the text information with the comprehender’s prior knowledge. Behavioral studies have shown that for narrative texts, the situation model contains information on aspects such as the where, when and why of the story (Zwann, Magliano & Graesser, 1995). In this study we used an inconsistency paradigm (Rinck, Hähnel & Becker, 2001) to investigate the functional neuroanatomy of situation model building. During the fMRI study, 20 students were scanned at 3 T while they listened to 32 short stories. Half of the stories contained inconsistencies with respect to the temporal or emotional aspects of the situation model. To avoid susceptibility artifacts in orbito-frontal and ventro-median cortex a spin-echo sequence was used. As expected, story comprehension elicited an extended, bilateral fronto-temporal network of activation. Using subtraction contrasts, we show that the integration of inconsistent emotional information was qualitatively different from the integration of temporal violations. Analyses of the time courses of activation are used for illustrating the regions’ different roles during language comprehension in context. Finally, correlational analyses suggest that right-hemispheric regions are activated in concert with their left-sided homologues.

The results are discussed with respect to neuropsychological theories of language comprehension. In particular, we focus on the contributions of the left prefrontal cortex and we would like to discuss in detail the role of the right hemisphere.

References:


Discourse psychologists assume that readers construct situation (mental) models as they attempt to comprehend texts in various genres, such as narrative and expository text. These constructive processes require the generation of inferences and connections among text constituents. A taxonomy of inferences has been proposed in the constructionist theory of Graesser, Singer, and Trabasso (1994) and in the event indexing model of Zwaan and Radvansky (1998), whereas Graesser, McNamara, and Louwerse (2003) have identified a taxonomy of cohesion signals and relations. Discourse psychologists are encouraged to consider a broad landscape of inferences and cohesion relations in their cognitive models and in their empirical investigations that collect behavioral or neuroscience data. We have recently been assembling components in computational linguistics (and what we might call computational discourse) that automate some of the mechanisms that are needed to build coherent situation models. Our most recent tool, called Coh-Metrix, analyzes texts on hundreds of dimensions of cohesion and language. Coh-Metrix incorporates lexicons, classifiers, transducers, syntactic parsers, shallow semantic interpreters, and latent semantic analysis. Coh-Metrix will undoubtedly offer greater precision in our analysis of language and discourse patterns in texts. This engineering feat will no doubt prove to be useful, but there are some pressing theoretical and empirical questions as to whether the particular computational modules have any correspondence to cognitive and brain mechanisms. Our objective is to identify some connections among theoretical claims, behavioral data, neuroscience, and computational linguistics/discourse.
Standardized psychometric assessments of reading comprehension such as Item Response Theory (IRT) are based on a one-dimensional ability scale of measurement for the purposes of convenient quantitative analysis (e.g., Lord, 1980). That is, a reader’s skill in understanding a collection of “psychometrically equivalent” texts is represented as a single number. However, this “single number” reflects a variety of items that are heterogeneous with respect to what a student needs to understand to answer correctly. For example, some items are directly answerable from the surface text; other items require interpretation of the text, a process that requires that a student use prior knowledge in answering. Modern theories of text comprehension (e.g., Cote, Goldman, & Saul, 1998; Goldman & Varma, 1995; Kintsch, 1998; Trabasso & Magliano, 1996; van den Broek, 1990; Zwaan, Magliano, & Graesser, 1995) assume that comprehension is a process whereby readers construct complex cognitive representations, known as “situation models.” Situation model components include information from the text, inferences based on the text, relevant prior knowledge, and inferences that relate the text and prior knowledge (Coté & Goldman, 1999; Kintsch, 1998). They can be represented as connected networks of propositions and relations among them. These directed graph structures specify particular patterns of knowledge relationships among propositions referenced in the text as well as those inferred by individual readers.

We are using a combination of discourse analytic strategies, multinomial logistic regression time-series analysis, and hidden Markov modeling methods to estimate the components of readers’ representations and thereby diagnose their comprehension strengths and weaknesses. Essentially, we use discourse analytic strategies to generate directed graphs for texts and then use the statistical methods to create probabilistic matches between proposition subsequences observed in readers free response summaries or answers to questions on the texts (Golden, 1998; Durbin, Earwood, & Golden, 2000). The directed graphs are populated by three “layers” of information: explicit text information, inferred from the text, and prior knowledge inferences. Connections among bits of information can be connected within layers and there can be inferred connections between layers. The pattern recognition process enables a diagnostic assessment of readers’ interpretations of text.

Figure 1 shows a graph structure of a situation model assumed to be associated with a subject’s understanding of a two sentence text. Figure 2 shows a connectionist model of response production. Assume that the subject/model generated the proposition sequence: M1, M3 in the production data. This sequential information is modelled as an activation pattern over the input nodes where input node M3 is most activated (since it was most recently mentioned) and input node M1 is partially activated (since it was mentioned prior to proposition M3). That is, the assumption is made that when a proposition is mentioned in the course of production, that node’s activation level decays gradually and the residual activation level can influence the system’s subsequent responses. The output proposition node M4 is activated
indicating that M4 is likely to be the next proposition mentioned by the subject/model. Activation levels and patterns are determined by types of connections in the situation model graph representation and determine the likelihood of which node will next be mentioned. In addition to describing the model in some detail, we will discuss its use in the context of a specific experimental study where 3rd, 5th, and 7th grade school children were asked open-ended questions about several narrative and informational texts. After illustrating the proposed methodology and demonstrating the types of results which have been obtained thus far, we will consider the relative strengths and weaknesses of our approach in comparison to other cognitive modeling methodologies. We will also discuss the implications of this methodology for advancing basic research in the field of text comprehension.

References


In this example, the original text was: “Mary wanted lunch. She checked her watch and bought some lunch.” Proposition M1 is a knowledge-inference while the text may be modeled as the proposition sequence: M2, M3, M4. The causal relation between M2 and M2 is called a Textbase-inference.
Figure 2. **Connectionist Computational Model.** The model implements the componential situation model described in Figure 1. Assume that the subject/model generated the proposition sequence: M1, M3 in the production data. This sequential information is modeled as an activation pattern over the input nodes. Specifically, input node M3 is most activated and input node M1 is partially activated. The output proposition node M4 is activated indicating that M4 is likely to be the next proposition mentioned by the subject/model.
How is negated information mentally represented during language comprehension? In language-comprehension research, negation has typically been investigated in the context of theories that assume that the representations involved in language comprehension are linguistic in nature (e.g., propositional representations). However, the last two decades have produced a great deal of evidence that comprehending a text is tantamount to the construction of a situation model, a mental representation of the state of affairs denoted by the text. Proponents of the standard view of situation models assume that the representational format of situation models is specific to the area of language processing (e.g., van Dijk & Kintsch, 1983). Recently, however, the notion that situation models are perceptual simulations of the referent situation has been gaining in importance (Barsalou 1999; Glenberg & Kaschak, 2002; Zwaan, Stanfield, & Yaxley, 2002).

Research conducted in the context of the perceptual-simulations view of language comprehension has until now focused on providing evidence that comprehenders construct perceptual simulations of the described situations when processing simple affirmative declarative sentences. As of yet, there is little research on the question of how particular linguistic operators, such as negation, are represented in perceptual simulations. Obviously, this problem needs to be tackled if the perceptual-simulations view is meant to hold for language processing in general. With respect to negation in particular, it needs to be clarified how negated text information can be captured in a nonlinguistic, perceptual representation.

The present work addresses the hypothesis that negated text information is implicitly represented in the sequence of processes that unfolds as the comprehender constructs a perceptual simulation for a negative sentence. More specifically, we hypothesize that a negation is a cue to the comprehender to do three things: first construct a mental simulation of the negated situation, second focus attention away from this representation, and third construct a mental simulation of the actual state of affairs (Kaup & Zwaan, in press; cf. Fauconnier, 1995; Langacker, 1991). The focus of the present work is on the first step.

In three experiments, we investigated whether the processing of a negated sentence leads to a perceptual simulation of the state of affairs negated by the sentence. All three experiments used the paradigm developed by Zwaan et al. (2002) for testing the perceptual-simulations view with affirmative sentences. As in Zwaan et al.’s study, participants were presented with sentences such as The ranger saw an eagle in the sky or The ranger saw an eagle in the nest, and afterwards saw a picture of the object mentioned in the verb phrase of the sentences. Participants judged as quickly as possible whether the object in the picture was mentioned in the sentence. For experimental trials, the correct response was always ‘yes’, but the picture either matched the implied shape of the object (outstretched wings for ...in the sky; folded wings for ...in the nest) or not (folded wings for ... in the sky; outstretched wings for... in the nest). Zwaan et al. found a strong match/mismatch effect. Response latencies
were significantly shorter when there was a match between the sentence and the picture with respect to the object’s shape than when there was a mismatch.

What can be predicted about negated sentences in this paradigm? If it is true that negation is a cue to the comprehender to construct a perceptual simulation of the negated states of affairs, then the negated sentences should yield similar match/mismatch effects to the affirmative sentences. Thus, if comprehending a sentence such as There was no eagle in the sky initially requires a simulation of an eagle in the sky, then this should be reflected in the response latencies elicited by pictures of an eagle with outstretched or folded wings, respectively. Latencies should be shorter if the picture matches the implied shape of the object in the situation that is being negated (i.e., outstretched wings) than when there is a mismatch (i.e., folded wings). Conversely, There was no eagle in the nest should lead to the reversed latency pattern. In this case, latencies should be shorter for a picture of an eagle with folded wings than for a picture with outstretched wings.

The results of the experiments support these predictions. In Experiment 1, participants were significantly (by approximately 30 ms) faster in the Match than in the Mismatch condition. Experiment 2 replicated this result, and ruled out the hypothesis that participants simply ignored the negative particles in the experimental sentences. In this experiment, participants were presented with comprehension questions that required processes that go beyond understanding the individual words in the sentences. The match/mismatch effect was highly significant even when only those participants with high accuracy on the comprehension questions (> 80%) were included in the analyses. Experiment 3 extended these results to other kinds of negative sentences. In addition to the ‘indefinite’ negations (e.g., There was no eagle in the sky.), participants were presented with ‘definite’ negations (e.g., The eagle was not in the sky.). The match/mismatch effect did not interact with definiteness but proved significant for both types of negations.

Overall, the results of the three experiments support the hypothesis that comprehenders routinely simulate negated situations when comprehending negative sentences. Further research is underway in our labs that directly addresses the second and third steps of the proposed three-step process, namely the shifting of attention away from the negated simulation and the construction of a mental simulation of the actual state of affairs.

References


Inference-making is a complex and necessary activity in the comprehension processes of daily life. Given its complexity, it has been explained from diverse theoretical viewpoints. At the same time, there has emerged a certain amount of contradictory data, and this poses some important questions. These concern the taxonomy of inferences, how and when inferences are processed and activated, and which of the theories best explains and predicts the creation of inferences.

One important factor that influences these questions is the previous world knowledge that gives rise to the production of inferences. Since the 1970´s, studies have examined the role of prior knowledge in comprehension (Anderson & Pitchert, 1978; Bransford & Johnson, 1972). Prior knowledge vitally affects everything we do, from perceiving and imagining, to text comprehension and creative problem-solving (Fincher-Kiefer, 1995). The empirical evidence for prior knowledge indicates that there is a monothematic relation between recall of new information and prior domain-specific knowledge. Specifically, it has been found that experts and novices differ with respect to access to relevant information and the sophistication of knowledge-based strategies (Artel, Shiefele & Scheider, 2001), and there are clear examples of how domain-specific knowledge influences inference-making related to this specific knowledge (León & Carretero, 1992; León & Pérez, 2001; León, 2003; Vonk & Noordman, 1992).

In this paper we focus on two different types of knowledge-based inferences: clinical diagnosis inferences and trait inferences. While a clinical diagnosis inference is a particular kind of inference that is produced in a specific knowledge domain (clinical psychology and psychiatry), trait inferences constitute concepts that represent the categorization of behavior using social and more general knowledge. Several questions are asked: Are the cognitive processes that participate in the generation of knowledge-based inferences the same though they use knowledge from different contexts (general or specific knowledge)? Given the critical role of previous knowledge, is the degree of expertise in a specific domain influential in accelerating knowledge-based inferences in different contexts (general vs specific knowledge inferences). Different types of data are presented indicating diverse possible explanations.
Sentence and Discourse Representation in the Two Cerebral Hemispheres

Debra L. Long

Co-authors:

Kathleen Baynes
Chantel Prat
University of California, Davis

Psycholinguists argue that readers construct and store in memory at least two interrelated representations when they comprehend a text: (1) a propositional representation containing the individual ideas (propositions) that are derived from each sentence and the relations among those ideas and (2) a discourse representation containing information about the context or situation to which a text refers. We have used a paradigm called “item priming in recognition” to examine how these representations are stored in the two cerebral hemispheres. In some studies, we have used the priming paradigm in combination with a lateralized visual-field procedure. We have found that the left hemisphere is very sensitive to the propositional structure of a passage. A concept from a passage is recognized faster when it is preceded by another concept from the same passage. Moreover, the amount of facilitation (propositional priming) bears a linear relation to the distance between prime and target concepts in the propositional structure of the passage. Facilitation is greatest when the target is preceded by a prime from the same proposition in the same sentence. This is true, however, only when targets are presented to the left hemisphere. In contrast, the right hemisphere is sensitive to the presence of concepts in the same passage, but appears insensitive to the propositional relations among them. One exception to this latter finding, the absence of propositional priming in the right hemisphere, is found among young adult readers who perform poorly on a standardized test of reading comprehension. They exhibit propositional priming in both cerebral hemispheres. We have also used the item-priming paradigm to study the representation of semantic information about passages. We have found that both hemispheres represent information about the context-appropriate senses of ambiguous words and this is true for both good and poor readers. Information about the themes and topics of passages also appears to be represented in both hemispheres, except in very good readers, who show representation only in their left hemispheres. Finally, we have used a variation of the paradigm to study sentence and discourse representation in brain-damaged populations including callosotomy patients and patients with left-hemisphere damage. We have found normal priming in the left hemispheres of callosotomy patients, but substantially diminished propositional priming in left-hemisphere damaged patients. We argue that the propositional representation resides primarily in the left hemisphere, whereas contextually appropriate semantic information is represented in both hemispheres.
Introduction

Recently, a considerable amount of research has been generated that examined the role of situation models in people’s comprehension and memory of texts. Presumably, the cognitive mechanisms that lead to the understanding and episodic memories for events operate independent of modality of experience. As such, while research on situation models in the context of text comprehension is important, it is also important to look at how this theoretical construct can be moved beyond language comprehension to provide theoretically interesting insights into cognitive processing more generally. Our presentation would provide a brief outline of our view of what situation models are and how they function followed by an in-depth exposition of how this concept can be applied to three areas outside of the area of language comprehension. In particular, these areas are inferring the goals of strangers, indexing situational continuity visually experienced event, such as narrative films, and virtual reality interactions, and autobiographical memory.

Situation Models

In our view, situation models are mental simulations of real or possible worlds. That is, these mental representations isomorphically capture the elements of a situation and relations among them that define the situation. There are a number of elements and relations that can be involved in the structure of the situation model. In providing this definition we distinguish between static and dynamic aspects of a situation model.

In terms of the static component, a situation is defined and bounded by a spatial-temporal framework. This is the region of space that contains the situation, and the stretch of time that that situation is in force. Within this spatial-temporal framework are tokens that represent entities. These entities can be things like people, animals, objects, abstract concepts, and so forth. Associated with these entities can be various properties. These properties can include external physical properties, such as size, color, or weight, and internal properties, such as emotional state, goals, and sanity level. Finally, there are structural relations among those entities within the framework. These can include things such as spatial relations, social relations, and ownership relations. The likelihood that properties and relations are included in the situation model is a result of the degree to which they play a functional role in defining the interaction among the critical situation elements. The more they are interacting or likely will interact, the greater the probability that they will be represented in the model.

In terms of the dynamic component, a series of spatial-temporal frameworks may be joined by a collection of linking relations that joined the frameworks in a sequence. These linking relations can be things like temporal and causal relations. These linking relations are grounded in the entities. This is because it is the entities that are moving through time, and which have causal interactions with one another.

Inferring the goals of Strangers

When we experience a narrative, we often encounter unfamiliar characters. Inferring the goals of these characters as they interact in the narrative world is critical
to understanding. According to the even indexing model, when inferring the goals of characters, readers should index a situation dimension, such as spatial setting, time, properties of the character, their actions, and the causal relationships between the events. This may also be true when we interact with strangers in the real world. Researchers at Northern Illinois University and Sandia National Laboratories have conducted studies to assess the extent to which individuals index situational dimensions to infer the goals of strangers. As is the case with narrative texts (e.g., Zwaan & Radvansky, 1998), individuals index multiple situational dimensions to achieve this goal. Moreover, it has been found that some dimensions carry more weight than others. For example, the actions and characteristics of a stranger have a greater predictive impact than do the spatial setting or the time in which the interaction occurs, or characteristics of the comprehender (e.g., physical and psychological states).

**Narrative Film and Virtual Reality**

Research testing the event indexing model in the context of narrative text has shown that readers monitor situational continuities, but that some dimensions are monitored more closely than others. Specifically, readers monitor changes in causal and temporal contiguity more closely than spatial contiguity (e.g., see Zwaan & Radvansky, 1998 for a review). The comprehension the visual medium provides a unique opportunity to assess the event indexing model to see if it can generalize to other kinds of experiences. Magliano, Miller, & Zwaan (2001) had participants view full-length feature films. Their study showed that viewers monitor temporal shifts more closely than narrative shifts in region or changes location of characters. These results indicate that the results obtained from narrative texts studies may generalize to other types of media. In a recent set of studies, we have looked at people’s performance in virtual reality situations. Of particular interest are flight simulations and ground combat situations. In these tasks, people are asked to interact in some virtual desktop environment. Afterwards, performance was coded with respect to the components as identified by existing situation model theory. An interesting new component is the inclusion of some representation of the self in this situation.

**Summary**

Situation model theory is a powerful way of looking not only at language comprehension, but cognition in general. We present here three different domains of thought where this theoretical approach has been successfully applied. It seems likely that there are other domains of cognition that are open to the theoretical insights and tools that have been successfully applied in the area of language comprehension.

**References**


Three Accounts of Comprehension Skill and Inference Making:
Knowledge versus Capacity and Suppression

Danielle S. McNamara  
University of Memphis  
Manuel de Vega  
Universidad de la Laguna  
Tenaha O'Reilly  
University of Memphis

Behavioral studies of reading comprehension have shown that skilled and less skilled comprehenders can be distinguished in terms of how often, when, and how they make inferences while reading. This paper reviews this evidence as well as three theories that have been proposed to account for these differences. One theory proposes that comprehension skills are determined by differences in working memory capacity, such that skilled readers are better able to make inferences while reading because they are able to hold in working memory more information from the text or discourse. A second explanation proposes that skilled readers are better able to suppress contextually irrelevant information. Evidence is reviewed favoring a third class of explanations which propose that skilled comprehenders more actively and efficiently use knowledge and strategies. In addition, neurological examinations of reading skill differences are reviewed and interpreted in terms of knowledge-based theories. It is argued that individual differences in comprehension skill depend on dynamics associated with knowledge activation during comprehension.
Pronoun resolution in a model for knowledge-based inferences.

Leo Noordman
Tilburg University
Wietske Vonk
Max Planck Institute for Psycholinguistics and Nijmegen University
Stefan Frank
Tilburg University and Nijmegen University
Mathieu Koppen
Nijmegen University

In a sentence like “Bob lost the match to Joe, because he was tired” the pronoun ‘he’ is ambiguous. Several factors may affect the resolution of the pronoun. One factor is foregrounding. On the basis of the fact that Bob is subject of the first clause and on the basis of the implicit causality of the verb in the first clause, Bob is foregrounded and ‘he’ refers to Bob. Another factor is world knowledge. Our knowledge about ‘losing’ and ‘being tired’ suggests that ‘he’ refers to Bob. A model is presented that simulates the process of knowledge-based pronoun resolution. The model deals with the problem of selecting the relevant world knowledge in the comprehension of narratives.

In the model, a story is considered as a sequence of story situations. A story situation is a combination of propositions that occur at one moment in the story. In the model, propositions and situations are represented distributively as points in a high-dimensional space. A story being a sequence of situations is represented as a trajectory through space. Statements in the story together with world knowledge affect the belief that unstated propositions can be the case at a particular moment in the story. Propositions that are likely to be the case have their belief increased, i.e., they are inferred and the story trajectory is adjusted. Consequently, the resulting trajectory matches knowledge better than the original story.

This model is able to simulate the processing of ambiguous pronouns. Possible instantiations of the pronoun lead to different propositions (‘Bob was tired’ vs. ‘Joe was tired’), corresponding to different points in the space. The process of pronoun resolution comes down to letting the vector representing ‘he was tired’ fall towards one of these points. The referent that is chosen and the time this takes depend on foregrounding and world knowledge. The results of the model agree with empirical data regarding the time course of pronoun resolution, reading times and error rates.
How ERP exposes integrative processes in text comprehension

Charles Perfetti
University of Pittsburgh

Co-authors:

Franz Schmalhofer
University of Osnabrueck

Michal Balass, Jessica Nelson, Chin-Lung Yang
University of Pittsburgh

How semantic integration occurs between two consecutive sentences has been an important research topic in the reading and text comprehension literature for a long time. As Haviland and Clark (1974) have pointed out, sentences contain both given information (what the reader knows already) and new information (what the reader does not yet know). A definite noun phrase of a consecutive sentence signals that this phrase can be mapped to a concept that has already been introduced and thus yields shorter reading times than without such an indication (Haviland & Clark, 1974; Experiment 1). According to the Given-New Strategy, a reader would attach the concept referenced by the definite noun phrase to the respective antecedent information of the previous sentence. For the sentences “Bill took out the picknick supplies. The beer was warm”, the connection between beer and picknick supplies would thus be cognitively established at the time when “beer” is read. There would be only a small difference to reading the second sentence after having read “Bill took out the beer” as a first sentence. The difference is, that the identical or similar concept is now already signalled by the orthography and phonology as opposed to only by an existing semantic relation. According to the lexical quality hypothesis of reading comprehension (Perfetti & Hart, in press) it can be assumed that in reading both orthographical and semantic relations will contribute distinct ERP-signals.

In research on predictive inferencing, it has been found that a faster reading time of a sentence can also be observed when some preceding sentence invites a predictive inference, even after a relatively long delay (Casteel, 2001), as in “No longer able to control his anger, the husband threw the delicate porcelain vase against the wall”. The faster reading time for a consecutive sentence like “He picked up the broken pieces of the vase” does not only indicate that the predictive inference had been drawn. More interestingly, it shows that it is still available at this later point in time and that it facilitates the processing of the related sentence. ERP-data can further inform such research (cf. van Berkum, Hagoort & Brown, 1999)

By collecting ERP-data, we explored in more detail, how referential bridges are established as a function of the different types of connections that may be established between a critical word and a concept of a preceding sentence. This concept may have been introduced by the identical word, by a paraphrase, or a respective event may have been previously inferred. In addition to these three experimental conditions, we employed a control condition, in which the critical concept had not been introduced at all, but where the sentences had used about the same content words.
Table 1 shows a sample text passage with its four versions for the four different experimental conditions. One hundred and twenty of such materials were constructed so that 30 materials could be used for each condition in a within-subject design.

<table>
<thead>
<tr>
<th>Explicit</th>
<th>Paraphrased</th>
<th>Causal Prediction</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>When Mildred, a Las Vegas salesgirl, carelessly dropped her cigarette near the woods, the result was a tragic fire. The fire engulfed many trees before being extinguished.</td>
<td>When Mildred, a Las Vegas salesgirl, carelessly dropped her cigarette near the woods, the result was tragic flames. The fire engulfed many trees before being extinguished.</td>
<td>When Mildred, a Las Vegas salesgirl, carelessly dropped her cigarette near the woods, the result was tragic. The fire engulfed many trees before being extinguished.</td>
<td>Mildred, a Las Vegas cigarette salesgirl, tragically crashed her car into the woods. The fire from her car engulfed the million dollars she had just stolen from the bank.</td>
</tr>
</tbody>
</table>

The 120 text passages were divided up into 4 sets of 30 passages each. The specific version of a passage (and thereby the specific experimental condition) in which a given set of 30 passages was presented was counterbalanced across eight participants. This implies that there were two subjects in each of the four groups. The two subjects of each group were presented with each passage in exactly the same experimental condition.

Eight students from an introductory psychology class from the University of Pittsburgh participated in individual sessions of the experiment. They were told that they would be reading several brief texts while their brain waves would be recorded by an EEG and that they should read the text in a careful manner because their comprehension would be tested.

The Experiment was run under the control of an E-Prime program. A rapid visual serial presentation (RVSP) technique was used for presenting each passage, word by word at the center of the monitor. A 15-inch CRT monitor working at 60 Hz refresh rate presented each word for 300 ms followed by an inter-stimulus-interval of 100 ms. Randomly distributed and on average after each forth trial, a comprehension question was presented after the reading of a passage, which the subjects answered by pressing one of two buttons. The subjects were given immediate feedback about the correctness of their answer.

A 128 Channel Geodesic Sensor Net (Electrical Geodesics Incorporated, Oregon, USA) recorded the EEG data. All impedances were kept below 40 K Ohm. Six eye channels allowed the rejection of trials with eye movements and blinks. ERPs were averaged off-line over the critical word for each set of stimulus materials in each of the four experimental conditions.

In analysing the data, we applied a subtraction logic between the experimental conditions. The differential effect of morpheme processing should be highlighted by the difference between reading the critical word (e.g. fire) in the paraphrased versus the explicit condition. The differential effect of establishing an intensional meaning representation should be highlighted by the difference between ERP-signals of the paraphrased and the inferencing condition. Finally, establishing an extensional discourse entity or situational unit should be highlighted by the difference between the inference and the control condition.
The results will be interpreted in terms of specific ERP-components, such as the N400, the P300 and the three lexical constituents which may be newly constructed or connected to already existing mental units when reading the critical word (Perfetti & Hart, in press). In addition, current theories of on-line inference generation (Schmalhofer, McDaniel & Keefe, 2002) will be applied for discussing the results.

References:


Mason and Just (2003) used fMRI to examine the cortical systems responsible for generating and integrating causal inferences; their results suggest that the bilateral dorso-lateral pre-frontal cortices are responsible for generating inferences, whereas right-hemisphere language areas are responsible for integrating these inferences into memory. In this chapter, we first develop a set of computational principles that are sufficient to account for these results, and then speculate about how these principles might be instantiated in the brain. Our goal is to develop a computational model that can be used to predict the outcome of both behavioral (e.g., sentence-reading times, memory performance, etc.) and brain-imaging text-processing experiments.
Time is an ubiquitous aspect of our life, and it also is an important piece of information derived from discourse. In several series of experiments, we investigated how readers represent temporal information in situation models created from discourse. Based on the event-indexing model proposed by Zwaan and his colleagues (see Zwaan, 1999) and the associative network model proposed by Bower and Rinck (2001), different aspects of the temporal dimension of situation models were investigated. In the experiments, we employed a variety of behavioral measures such as reading times, probe reaction times, eye movements and fixations, recognition rates, and consistency ratings.

In one series of experiments, we (Rinck, Hähnel, & Becker, 2001) used the inconsistency paradigm to show that temporal information is represented in situation models and monitored during reading, just as other dimensions are. With self-paced sentence-by-sentence reading, readers showed increased reading times when they encountered temporal information which was inconsistent with the temporal aspects of their current situation model. In related experiments, we (Rinck, Gámez, Díaz, & de Vega, in press) measured eye movements and fixations during the unrestricted reading of complete texts, finding that temporal inconsistencies caused regressions towards related temporal information. Two other experiments addressed the potential confound of temporal and causal relations (Rinck, 2002). The temporal order of events may be important only because of cause-effect relations: earlier events may cause later ones, and not vice versa. However, the experiments revealed that temporal relations are being monitored even when they are not causally relevant.

In second series of experiments, we (Rinck & Bower, 2000) investigated how temporal distance interacts with spatial distance represented in situation models. We found that story time distance (i.e., fictitious time passing by description within the narrative world) and spatial distance (i.e., described distance between entities within the narrative world) affected accessibility of situation model entities additively. Discourse time distance (time passing in the real world during reading), however, did not affect accessibility.

In a third series of experiments, we systematically compared the temporal dimension of situation models to other dimensions. In two experiments, we (Rinck & Weber, 2002) experimentally varied the temporal, personal, and spatial continuity of sentences, complementing the regressional approach employed by Zwaan and his colleagues (see Zwaan, 1999). We found that for each dimension, discontinuities caused reading time increases, with slightly larger effects for the temporal and personal dimensions. In two other experiments, we again employed the inconsistency paradigm to compare temporal, spatial, emotional, causal, and goal-related inconsistencies to each other (Rinck & Hähnel, 2002). In line with earlier studies, the strongest effects occurred for emotional and temporal inconsistencies, and the weakest effects for spatial ones.

The results of these experiments converge on the conclusion that time is an important dimension of situation models, which is routinely monitored and represented during discourse comprehension. In my talk, I will discuss the relevance of these data for theories of discourse comprehension, and I will give a critical evaluation of the behavioral measures used in the described experiments. Additional experiments were conducted by Evelyn Ferstl and her colleagues, using fMRI to
assess the neuropsychological basis of processing of temporal and emotional inconsistencies. These experiments will be presented by Evelyn Ferstl.

References:


Many of the experimental findings about inferencing and other higher level language processes have been explained within general cognitive frameworks. One of these frameworks (Kintsch, 1998) assumes that more knowledge becomes initially constructed than what becomes more durably represented in the mind after a context-sensitive integration process has selected the most appropriate meaning. Many computational models that were developed within such frameworks succeeded in accounting for a large variety of experimental findings (Singer & Kintsch, 2001). In my presentation, I want to explore whether such computational models can also account for results from neuroscience, such as fMRI- and ERP-data at some level of detail.

I will first describe the KIWi-model (Schmalhofer, 1998). This model assumes that three levels of representation are formed, verbatim, propositional and situational representations. Whereas the intensional meaning of a text is represented by propositions, its referential situation model is encoded by perceptual symbols (Barsalou, 2000; Zwaan, in press). How the model works will be shown by having it predict several results from behavioural experiments. Thus, it will be described how the recognition and verification of sentences are performed. Differences in the encoding between low and high knowledge readers will similarly be accounted for. And it will be derived how the model explains the priming effects in the pronunciation of inference related words (McDaniel et al., 2001).

For applying the model to neuroscience data more fine-grained assumptions about the localization and timing properties are needed. More specifically, empirical results and respective theoretical assumptions from Long & Baynes (2002) and Mason & Just (in press) will be used to derive predictions for an ERP-experiment conducted by Perfetti et al. (this conference) and an experiment where a proposition verification task was performed under hemi-field presentation restrictions. It is argued that it is indeed possible to enhance and refine existing computational models that have been developed to account for behavioural data with assumptions when and where such higher level language processes occur in the human brain.
Advances in the study of text comprehension in general and inference processes in particular have gone hand in hand. One focal point of comprehension theory has been the automaticity of text inference processes. In this study, Jacoby's (1991) process dissociation (PD) procedure was applied to the diagnosis of the automaticity of text-inference retrieval. The emphasis was on bridging inferences. Bridging inferences contribute to text coherence by identifying the connections among ideas, whereas elaborative inferences simply specify sensible extrapolations from text. Prior studies have shown that bridging inferences are indistinguishable from explicit text ideas on numerous measures, suggesting similar long-term memory (LTM) representations for the two; whereas elaborative inferences are inferior. We evaluated the LTM representations of explicit and implicit text ideas using the extended process dissociation procedure (Buchner, Erdfelder, & Vaterrodt-Plunnecke, 1995). Three experiments used the three phases of the Jacoby's process dissociation experimental paradigm to partition the controlled, recollective contributions to text retrieval from the automatic, familiarity-based contributions. The experiments showed that (a) explicit text ideas are more strongly supported by both controlled and automatic influences than are inferences, (b) support for the recognition of inferences is predominantly controlled, and (c) there may be a modest automatic contributions to the retrieval of bridging inferences but not elaborative inferences. These results diagnose informative differences between the LTM representation of explicit text ideas and text inferences. They bear on issues including (a) whether both recollection and familiarity influence text retrieval; (b) the automaticity of the encoding of text inferences, and (c) what it means to incorrectly recognize the implication of a text.

References

Assessing Index Dominance During Situation Model Construction

David J. Therriault
Florida State University

Co-authors:

Mike Rinck
Dresden University of Technology

Rolf A. Zwaan
Florida State University

An important goal in discourse research is identifying what elements of text are represented in memory during reading. Prior research provides evidence that readers create different levels of representation (i.e., surface, textbase, and situation model) (van Dijk & Kintsch, 1983; Fletcher & Chrysler, 1990; Kintsch, Welsch, Schmalhofer, & Zimny, 1990). Our focus is on situation-model representations. We conceptualize a situation model as a complex mental representation that includes information from the text as well as from the reader’s background knowledge. According to the Event-Indexing model (Zwaan, Langston, & Graesser’s 1995), readers are sensitive to multiple critical dimensions when constructing situation models. These dimensions are: space, time, protagonist, causality, and intentionality.

A host of experimental studies have examined single dimensions of situation models (for a review, see Zwaan & Radvansky, 1998). In contrast, few attempts have been made to gauge the relative contribution of several dimensions simultaneously (Scott Rich & Taylor, 2000; Rinck & Weber, submitted). One stumbling point has been the inherent difficulty in equating dimensions along a similar metric. It is unclear if a single metric could be used to equate a shift in time (e.g., an hour later) with a shift in space (e.g., changing a scene from Chicago to Tallahassee). In response to this difficulty, we have developed an inductive method that can be used to compare the relative strength of situation-model dimensions simultaneously.

Two experiments were conducted examining the dominance of protagonist, space, and time dimensions in narrative situation-model construction. Time and protagonists terms are ubiquitous in situation model construction. Time is necessary to understand cause-and-effect and protagonists are necessary to convey meaning about focal entities. The importance of space in situation-model construction remains an open question. Zwaan and Oostendorp (1993) and Hakala (1999) have argued that readers do not normally monitor the space dimension unless specific task demands are imposed. Consequently, we predicted that protagonist and time dimensions would be more dominant than space in our experiments.

Participants read narratives and were specifically instructed to attend to a single dimension (time, protagonist, or space) while their sentence reading times were recorded. Questions, presented at the end of each passage, were yoked to instructions to further motivate participants to attend to a single dimension. For example, participants who received instructions to focus on the time dimension answered time questions exclusively. Participants who received protagonist-focused instructions answered questions about characteristics of the protagonist exclusively. Critical-sentence reading times were then analyzed for shifts in dimensions highlighted or not highlighted by the instructions. Increases in these readings times demonstrate the importance of the (non-focused) dimensions in situation model construction and their imperviousness to task demands.
Results indicated that shifts in time and protagonist dimensions always increased reading times, irrespective of the instructional condition. Shifts in space, however, only increased reading times if the instructional conditions focused on space. The results support our prediction that time and protagonist dimensions are more critical than space (i.e., time and protagonist cannot be ignored) and appear to be impervious to the task demands we imposed. In our talk, we will discuss the functionality of our methodology and the implications of the data with respect to discourse comprehension theories.

References


Hemisphere differences in the processing of negative and positive emotional inferences

Isabelle Tapiero

&

Virginie Fillon

University of Lyon 2 (France)

A lexical decision task combined with a visual split field methodology were used to study hemisphere differences in the generation of emotional inferences. Subjects read short passages that implied a character’s emotional state, either with a negative valence (e.g., guilt) or with a positive valence (e.g., happiness). Importantly, the character’s emotion was implied in the text and was never explicit (adapted from the materials used in Gernsbacher, & al., 1992, Experiment 2). Although the whole passage implied an emotional state, the last sentence of each passage was intended to induce readers to generate an inference in relation with the character’s emotional state implied. For each passage, the inference was tested with a lexical decision task that immediately followed the offset of the final word of the last sentence. The probe words either referred to the emotion that matched the textual information (e.g., guilt/matching emotion word), or to an emotion with an opposite valence (e.g., pride/mismatching emotion word). The probes items used in the lexical decision task (words and non words) were presented in the left and right visual fields. Predictions were that greater priming effects should be observed in the left visual field (Right Hemisphere) when the probes represent mismatching emotional inference. Indeed, it has been shown that when people view words in the left visual field (Right Hemisphere), they respond fastest to test words related to less frequent or contextually inappropriate meanings of ambiguous words (Burgess & Simpson, 1988). In the right visual field (Left Hemisphere), only matching emotional inferences should be activated because the left hemisphere is dominant for contextually relevant information. In addition, priming effects should be stronger in the left visual field (Right Hemisphere) when the emotional state implied by the text is negative, whereas the opposite should occur when a positive emotional state is implied by the text. The main results show evidence for the existence of hemisphere differences in the generation of emotional inferences and indicate that emotion has a critical role in the construction of a coherent situation model.
A central component of successful reading comprehension is the construction of a coherent representation of the textual information in memory. The process of constructing such a representation involves the generation of inferences that connect the various parts of the text to each other and to readers’ semantic knowledge. In this presentation, I will discuss a theoretical model that integrates the various subprocesses that have been identified in the literature and pursue empirical implications.

The presentation will consist of four components. In the first section I will review prior behavioral data on the inferential processes during comprehension and on the relation between these processes and the eventual memory representation. In particular I will describe research on different types of inferences (e.g., backward, forward, to prior text, to background knowledge) and the circumstances in which each are made; I will also discuss the psychological properties of the resulting representations, in terms of retrieval and accessibility in memory and perceived importance. In the second section I will present a computational model that integrates these processes. This model, the Landscape model, assumes that over the course of reading, concepts and propositions fluctuate in their activation as a function of the text itself, of the inferential processes described in the first section, and of the automatic spread of activation through background knowledge and the developing memory representation. Concepts and propositions that are co-activated during the same processing cycles become connected in memory, with the strength of their connections a function of their individual levels of activation. At subsequent cycles these connections can be strengthened. Thus, an episodic memory representation gradually evolves, with the fluctuating activations in each processing cycle providing continual updates. Thus, on-line activations and off-line representations are connected in a dynamic fashion.

In the third section, I will summarize several studies that test unique properties and predictions by the Landscape model. In some cases, this involves reanalysis of data already in the research literature (e.g., studies on resonance by O’Brien & Myers); in other cases, this involves data collected with the explicit purpose of testing the model. In addition, I will discuss how the model may capture the processes of different populations of readers. In the fourth section, I will briefly review data obtained from neuroimaging techniques. At present, we have conducted several studies on hemispheric differences in the way that different types of inferences are processes. By the end of Spring we hope to have supplemented this data with results from several fMRI studies. Both sets of data will be discussed in the context of the Landscape model.

I will conclude by some remarks on what the combination of computational and neurological data has taught us about the cognitive processes in discourse processes.
An experiential framework of language comprehension

Rolf Zwaan
Florida State University

A framework for the comprehension of words, sentences, and discourse is proposed that is grounded in perception and action. The basic premise is that language comprehension is the mental simulation of the referential situation. These mental simulations are constrained by the linguistic input, the processing capacity of the human brain, and the nature of human interaction with the world. Words diffusely activate experiential traces. During construal, traces activated by different words in a clause or sentence mutually constrain each other to yield a simulation of experience. Successive construals are continuously integrated to yield a mental simulation of a series of events. The relative ease of this integration process is affected by, among other things, (a) the extent to which the transition among construals is consistent with human experience, (b) the degree of overlap between the current construal and the content of working memory. It is shown how this experiential framework can account for a range of empirical findings that traditional amodal propositional models cannot, while still being able to account for the findings that amodal systems have produced.